

nature

24 November 1977

No jobs left for the boys

A CRISIS is approaching in the employment of scientists in the academic and closely-allied worlds of the UK. Nothing explosive will herald its arrival; there will be no dramatic day on which we shall all wake up to find the newspapers full of it. Rather, over the next few years it is going to be more and more difficult for young people of an academic frame of mind to find a satisfying career. And no-one needs reminding that the influx of young staff is essential to the continued vitality of most academic departments.

The reasons for this crisis have often been discussed in these columns; they stem from three clearly defined factors. First, relatively few people are retiring from university posts at present, or, for that matter, from research-council laboratories or the scientific civil service. The enormous university expansion in the middle 1960s has left the system with large numbers of people now middle-aged: fewer than 1% of all university staff retire this year. The civil service has a somewhat more uniform age distribution, but even so less than 2% retire annually at present.

Second, for the past four years university finances have not allowed vice-chancellors the luxury of filling even such posts as have become vacant without asking the most searching questions about the prudence of taking on a new long-term commitment. The resultant partial freeze on hiring has nudged the student-staff ratio up to around 10:1, or from another point of view has called for an increase in the productivity of university staff.

Third, long term projections of student numbers unambiguously point to smaller universities in the 1990s. For the past ten years the decline in the numbers of those of university age (following the post-war baby boom) has been offset by a steady increase in the percentage of these who seek, and are qualified to receive, higher education. This percentage now seems to have peaked and to be in slight decline whilst the numbers in the relevant age group are rising to a new peak in 1981. For the past ten years then and the next five it seems as if substantial demographic swings will have been moderated by changes in the perceived desirability of going to university, resulting in a slow and controllable expansion in student numbers. But what happens beyond 1981 is anybody's guess. One thing is certain: that by the early 1990s numbers in the relevant age group will be more than a quarter down on those for the early 1980s. So the problem of adjusting university staffing to cope with this demographic change is a severe

one now, and will presumably become even more severe in the late 1980s as those educated at the time of the peak start to look for academic posts.

Thus, to get a job in 1977, you need to find someone who is retiring from a university that is not plagued by financial problems and is confident that in the long-term its student numbers will hold up. Needless to say you will not be the only applicant. And you will find the same competitors pursuing the same limited opportunities in research-council establishments. A supply of suitable people in excess of the demand for them must be one of the factors ensuring that low academic salaries hardly pose a headache to the government, despite mass lobbies and palpable inequities.

How do we get ourselves out of this bind, ensuring that bright minds are not wasted? Clearly some specific initiatives, such as support of more young scientists by the Royal Society, help to take a little of the strain. But it is unlikely that the problem will be seriously alleviated without some radical steps being taken—and in present circumstances these steps can hardly be vast expenditures on new laboratories. If they are we run the danger of acquiring yet another group which grows old together and doesn't start to open up further employment opportunities until the year 2010.

The unspectacular alternative (it can hardly be called a solution) is for very serious examination of the knotty problems of tenure and retirement. In the past ten years scientists, reflecting a national trend, have become ever more cautious of any sort of employment which left them without security above the age of 30, and have then felt safe until the age of 65. The civil service have, of course, led the way in providing a lifetime of stability. There is certainly a lot of good in buffering scientists from short-term change, but is the only alternative a thirty-five year guarantee of employment or, put another way, exclusion for thirty-five years of someone else from the same employment? Surely not when times are becoming progressively harder; some new form of contract of employment must be worked out which can respond to the aspirations of young scientists for employment without humiliating senior colleagues who have given a lifetime of service. Talks are already proceeding on these matters between various interested parties—those involved need every encouragement to come up with some flexible proposals as soon as possible. □



Will the voice of science of the Third World be heard?

Michael J. Moravcsik of the Institute of Theoretical Science, University of Oregon, appeals to third world scientists

MORE than nine tenths of the scientific activity of the world takes place in countries comprising one quarter of the world's population, and most of the Third World at the present is virtually 'out in the cold'. The result is a strong inferiority complex, a necessity of subsisting on scientific and technological 'hand-outs' from the advanced countries, and a society which has not felt the impact of the dynamic world view that the practice of science can engender.

Some of the reasons for this state of affairs are internal to the developing countries themselves, some external. Much of science development must be done by indigenous manpower, with local resources. Other aspects of such a development process hinge on cooperation from the outside since science is a collective undertaking.

It is also a somewhat esoteric undertaking. Its aims, results, and needs do not automatically percolate into society which lacks the background and perspective to initiate interaction with science. Therefore scientists themselves must establish communication with people outside their ranks, and attract support for science making.

In this respect, again with some notable exceptions, the small scientific communities of the developing countries have been quite ineffectual in the past. There are two main groups who should have heard their voice: the policy makers in their own countries, and the world scientific community.

It is rare to encounter a developing country where decisions affecting the provisions for local science are significantly influenced by the active, working scientists of the country, instead of by civil service bureaucrats or by one-time scientists of advanced age who have long lost whatever contact and involvement they might once have had with actual science making. Such 'science policies' therefore are often restricted to the discussion of vague generalities, and in any case deal predominantly with technology and not science.

It is even rarer, however, to encounter a forceful and articulate voice from the Third World speaking out to the worldwide scientific community on behalf of the second class members of this community who happen to live in the developing countries. The large, international magazines used for communication among scientists seldom contain analyses and proposals for reform from representatives of the Third World and the large international conferences very rarely hear from such representatives about ways to make the practice of science more equitable throughout the world. The problem lies outside the consciousness of the overwhelming majority of working scientists.

Opportunity: a UN Conference

The next two years appear to offer a favourable opportunity to effect a change in this state of affairs. Sometime in 1979 a huge United Nations World Conference on Science and Technology in Development will take place at which it should be possible for science in the Third World to speak out, should it decide to do so.

In particular, national position papers are being prepared, which discuss a country's needs and cooperative capabilities and put forward specific ideas and proposals with respect to any of the four main areas of the conference; science and technology for development, institutions, the United Nations system, and science and technology and the future.

It has been urged that the preparations of these country papers throughout the world be done in full consultation with all interested parties. Yet, it is quite possible that the

preparation of these most important documents in many countries is being done by a small group of people who are perhaps not even involved in scientific activities.

An appeal

Thus the first appeal I would like to make is that scientists in developing countries make a most strenuous effort to partake in their own countries in the formulation of the country papers. There is no time to be wasted in doing so: the country papers are being written right now, and thus the input must be forthcoming promptly and forcefully. There are perhaps three main relevant areas.

- First, in the area of scientific education. Especially in the initial stages of the scientific development of a country, at least some of the indigenous scientific personnel must be educated abroad, and in this process cooperation from the scientifically advanced countries is an absolute necessity. The problems that arise are not only financial but also centre around the special needs such 'foreign' students have in the context of the education they receive abroad. Proposals have been formulated previously in a profusion of articles and meetings. Action is now needed.

- The second area of concern should be scientific communication. At the present all scientific communication channels are structured so that the scientifically rich get richer, and the scientifically poor poorer. Journal distribution, report and preprint circulation, personal visits, study leaves, conferences, and so on are all designed to yield scientific results tomorrow and hence neglect the yet evolving scientific institutions which promise to produce only the day after tomorrow. This subject has also been fully discussed in the past.

- Finally, and perhaps most importantly, one must create a worldwide atmosphere of opinion in which it is recognised that partaking in the practice of science is a legitimate, nay crucial, ingredient in the overall development process for all countries of the world. The condescending view that doing science is for the rich, and that the developing countries should rather strive to become better fed intellectual and technological slaves to the advanced countries must be changed. Development is a broad concept, and it includes equitability not only in GNP per capita, but also in the self-image of countries with respect to their overall contribution to and standing in humanity. Science development is a very significant element in this.

The second appeal I want to make to scientists in the Third World is to urge them towards articulation through an alternate channel. While much of the formal activity of the conference will be through the process of country papers, the atmosphere of the conference, the stand of individual delegates and speakers, the structuring of the discussion, and in fact even the details of the various country papers from the scientifically more advanced countries can all be significantly influenced by forceful communication through direct channels. Thus it seems that scientists from the Third World should 'campaign' by writing and speaking to fellow scientists the world over, to professional societies and journals, and to the conference organisers, the special office for the conference, headed by Joao F. de Costa, at the United Nations in New York. The more specific and programme-oriented statements and communication are, the more likely they will be assimilated into some action.

In summary, this appears to be the time for scientists of the Third World to rise and let their voice be heard. The psychological circumstances are favourable because of the huge build-up the conference is getting in political terms. The channels are also available. All we need now is action.

□

Combatting rickets

Researchers, nutritionists and policy-makers face a problem over rickets in Britain, as **Alastair Hay** explains

A QUICK survey of the British public would suggest that its nutritional problems are not due to a lack of food. Yet even in the midst of overconsumption a specific nutritional deficiency—that of vitamin D—has been defined in certain sections of the population. Rickets, known to generations of Europeans as the 'English disease' is now a problem in the children of Britain's Asian community and occurs sporadically in some of the infants of the country's poorest populations. It was first reported as being a serious problem in the Asian children of Glasgow almost sixteen years ago and has since been observed in one Asian community after another throughout the country.

The summer issue of *Science for People*—the publication of the British Society for Social Responsibility in Science (BSSRS)—asserts that the UK Department of Health and Social Security (DHSS) actually causes rickets. The journal is noted for a blunt approach to the issues it tackles, but this attack will not be viewed altogether unfavourably by many clinicians concerned with the health of Britain's Asian population and particularly with the problem of rickets. They point out that the problem is well known—perhaps too well known—and that there has been no shortage of surveys and quotable figures. They claim that most of the answers to this proven dietary deficiency are known, but that the DHSS has so far failed to adopt a positive approach.

Faced with such hostility the DHSS has been giving the matter serious attention. The department's Committee on Medical Aspects of Food Policy (COMA) is reviewing the whole question of bone disease contracted as a result of nutritional deficiency. The committee's deliberations have been somewhat protracted and are still by no means completed. But a report is expected that will produce some recommendations for action.

It is children who suffer from rickets; the same bone disorder in adults is termed osteomalacia. It results from an inadequate supply of vitamin D. The vitamin is required for the active absorption of calcium and phosphorus from the gut. When this mineral supply is reduced in vitamin D deficiency, the body, in order to maintain blood mineral levels, has recourse to its only other source of supply, the bones. Bone growth ceases, the skeletal structure become progressively weaker

and eventually deteriorates to produce 'bow-leg' or 'knock-knee' rickets.

Although biochemists are having a field day trying to elucidate vitamin D's exact mode of action, at the nutritional level the problem would seem to be simply that of ensuring an adequate supply. But the solution is not as easy as it seems, which could be one explanation why the DHSS is taking so long to formulate a policy. The question that needs answering concerns the most effective vehicle for administering the vitamin: should it be done through chapati flour, or milk, or vitamin D capsules, or tablets?

As a major constituent of the diet



of most Asian families, chapati flour would seem to be an ideal medium through which to add extra vitamin D to the diet. The technical problems of mixing the vitamin and of its preservation have been solved, and clinical trials using vitamin D-fortified chapati flour have shown its effectiveness as a method of administering the vitamin. A problem may arise with regard to the concentration gradient. It is children, not adults, who are most in need of extra vitamin D, and adults consume more flour per head. The answer would be to have two types of flour, one fortified and recommended for children, the other unfortified. But as a spokesman for a bread company says: "Someone has to pay for the process".

Milk poses different problems. There are no legal constraints to forbid fortifying chapati or brown flour with nutrients; this is not true of milk. Britain is bound by an EEC directive forbidding any alteration to the composition of milk. When the issue was discussed in Brussels, several European countries were in favour of legislating for a 3½% fat level. Britain fought this, primarily on financial grounds, and insisted on the right to have milk virtually unaltered from cow to con-

sumer (most milk would have to have fat removed to attain the 3½% figure). The British case was conceded, but at the time Asian rickets was not considered to be a problem; now it is. Before embarking on any milk fortification programme Britain would first have to return to the negotiating table, which could be embarrassing, but little choice may remain if the DHSS comes out in favour of using vitamin D-fortified milk. The arguments in favour of its use are strong. As children drink more milk than adults, it would reach the population most in need of treatment. Those in favour of milk point out that Canada fortifies milk, milk products and margarine with vitamin D; the United States does the same and much else besides.

As for the third suggestion that vitamin D capsules or tablets be dispensed either at clinics or in schools, staff in some schools are opposed on the grounds that the procedure of identifying children in need from the school register is too complicated.

As breast-fed babies in Britain rarely develop rickets any health education policy ought to discuss the merits of breast feeding as well. The reason for this immunity is not yet clear, but is probably related to the fact that the breast milk of vitamin D-replete mothers contains a fairly high concentration of the vitamin present as vitamin D sulphate. Some groups of immigrant women stop breast feeding when they arrive in the country and resort to doorstep milk.

No significant impact

It is generally acknowledged that the health education programmes dealing with rickets have failed to make a significant impact in the Asian community. Some clinicians argue that these programmes have failed dismally in alerting local doctors and health workers to the seriousness of the problem. They add that there is great inertia at the level of the community physician, that many doctors are reluctant to become involved in preventative campaigns, and that many health workers have not been informed that the problem exists.

That rickets is present in the Asian community, and a serious problem, is not in doubt. Some may query the figures for the prevalence of the disease, and argue for more information. Others will say the evidence is now so overwhelming that action must be taken. The DHSS recognises that many doctors expect its COMA committee to make some positive recommendations stressing the need for a food fortification programme backed up by a good health education policy reaching all the Asian community. □

More cash for British science

THE UK Advisory Board for the Research Councils (ABRC) is considering how to divide the second windfall to have landed in its lap this month amongst the five research councils. In reply to a question in the House of Commons late last week Shirley Williams, Secretary of State for Education and Science, announced an extra £4 million for the science budget which is to be spent on "new capital work" in 1978-79. It is in addition to the £4 million allocated to the research councils in the mini budget at the beginning of the month.

The new £4 million is to come out of the £400 million allocated to the construction industry in the mini budget. It is strictly for building and, because of its small size, is more likely to be used for improving or extending existing facilities than starting up new ones.

Professor Geoffrey Allen, Chairman of the Science Research Council, which is almost certain to get the lion's share, would like to see the money going towards finishing construction already begun and extending central user facilities. He feels that this might be an opportunity to complete the nuclear structure facility at the Daresbury Laboratory and extend facilities for the synchrotron source there. The laser facility and spallation neutron source at the Rutherford Laboratory might also come in for a share of the cash.

Another priority in the SRC is likely to be helping universities provide facilities for marine technology and polymer engineering. These are research areas considered to be of national importance and currently receiving special treatment from the SRC. The

only new project which might be considered is an electron beam lithography unit for microelectronics which could be set up in existing buildings at the Rutherford.

The cost of implementing these projects alone could come to £3 million. When the cost of the bids for support from all the research councils is added up, it will undoubtedly be much more than £4 million, so the ABRC is in for a difficult task. Nevertheless, it hopes to have made the necessary decisions by the end of next week or at the latest by its next formal meeting which is set for 9 December. A decision on the earlier £4 million, however, is unlikely to be made public before the end of the year although it has been made and is awaiting Mrs Williams' approval. It will probably simply be added on to the total science vote for 1978-79 which will be announced sometime in the new year.

Judy Redfearn

Leading Soviet dissidents in London

DURING the last month, three leading figures from the Soviet dissidence and human rights movement have visited London. According to the most recent arrival, Dr Khronid Lyubarskii, an astrophysicist, their presence in the West reflects a determined effort on the part of the Soviet authorities to overthrow all opposition, neutralising the leaders by removing them from the scene. Neither Dr Lyubarskii nor Dr Valentin Turchin, the cyberneticist, who were expelled together on 14 October, had any real desire to leave the Soviet Union. They had applied for emigration only when the authorities had made it clear that they must choose between foreign exile or a protracted stay in a labour camp. The third visitor, Professor Mark Azbel, a theoretical physicist, had indeed had a genuine wish to leave the Soviet Union, having applied several years ago for a visa for Israel. After several years as a *refusnik*, Azbel finally received his visa last June.

It is no coincidence that all three are scientists. The dissident movement has a strong bias towards the scientific professions, since, in Dr Lyubarskii's words, "a scientist works in information—it is therefore natural that he begins to think about the society around him. It is the critical mind of scientists that leads them to think of human rights".

Their record of service to the human rights movement is impressive. As early as 1970 Turchin, together with Roy Medvedev, signed the second Sakharov letter. In 1972, Lyubarskii

was sentenced to five years in a strict-regime Labour camp for his human rights activities, and while there organised the observance of an annual Political Prisoners' Day (30 October), to be marked with protests and hunger strikes. For this further transgression, Lyubarskii was transferred to the Vladimir prison for the duration of his term. Both Turchin and Lyubarskii became members of the illicit Moscow group of Amnesty International, and were associated with the Helsinki monitoring group.

On the arrest of Aleksandr Ginsburg last April, Lyubarskii took over the administration of the Solzhenitsyn fund for aid to political prisoners. Unlike many Jewish activists, Dr Azbel feels it better to keep the Jewish emigration movement separate from the dissident movement for reform and human rights. This is not a distinction, however, which the Soviet authorities accept—application for emigration is an antisocial act, just as much as signing a petition of protest, and is punished in the same way by dismissal from one's job, with no possibility of continuing one's professional career.

Dr Lyubarskii gave an interesting estimate of the numbers of scientists penalised for their activities:

involuntary foreign exile	a few
arrests	tens
expelled from jobs	hundreds
routine restrictions after	thousands
signing protests etc (refusal of	
foreign passport,	
administrative restrictions).	

It should be remembered that punitive dismissal is far more serious a penalty than loss of job. In the Soviet Union no academic article can appear without a certificate from the academic institution where the author is employed. No 'freelance' submissions are permitted. "If Einstein or Faraday had lived in the Soviet Union, they could never have published their theories" stressed Azbel.

How well a scientist can survive professionally under such conditions seems to depend very largely on his particular field. In spite of several years without employment or professional facilities both Turchin and Azbel have found themselves able to continue their professional careers in the USA and Israel respectively. For Dr Lyubarskii, however, the outlook is bleak. His special field, the physics of planets, meteors and meteorites has moved forwards at such a pace that he feels doubtful whether he will ever catch up after almost six years with no access to journals.

Whatever their new professional commitments, all three are firmly committed to campaign for fellow-scientists still suffering constraint and imprisonment, and, indeed, they came to England for this purpose. They will campaign for physicist Yuri Orlov—now in prison; biologist Sergei Kovalev—in a labour camp; mathematician Viktor Brailovskii, Azbel's successor as leader of the Sunday seminar, who is threatened with a treason charge; and the young mathematician Anatolii Shcharanskii whose trial on treason charges now appears to be imminent.

Vera Rich

Soviets honour teamwork and technology

It is sixty years after the revolution—but the 1977 Soviet State Prizes for Science and Technology show no sign of anniversary euphoria. The works cited seem genuinely to reflect the current state of progress, without any attempt to honour 'great names' on this special occasion. There is considerable emphasis on teamwork: the majority of citations involve long lists of project-directors, assistants, and technicians, so that the announcement of 38 awards with the names, academic or professional titles, and affiliations of the recipients fills a page of *Pravda*.

The citations span the usual wide range from mathematics to biology, with occasional incursions into fields not normally considered sciences in the West—Marxism and its ancillaries of history and economics. Current Soviet doctrine, as expounded at recent Party Congresses, demands that science should serve the national economy. Accordingly, several prizes go to research directly related to official priorities—a set of geological maps of Western Siberia, special equipment for land reclamation, fisheries equipment, and the use of discrete transformations in automatic control. The recent voyage of the icebreaker *Arktika* to the North Pole won a State Prize for her designers and engineers (though not for the scientific team aboard).

The medical sciences, which have recently received a new priority including the establishment of the title of merit "People's Doctor of the USSR", are well represented. Three science prizes go to long-term medical research: two—for haematology and prophylactic helminthology—mention "cycles of works" published over the last 15 years, while the third—one of the very few individual awards—cites a haematological monograph which also implies a long-term basis of research. Medicine also receives two technology prizes, where the citations seem rather more adventurous—new equipment for hyperbaric oxidation, and the "clinical development and implementation" of methods of limb transplantation.

In a year of setbacks for the space programme, it is not surprising to find space ignored in the technology section, although it does receive one science award—for research into the X-ray emission of the sun.

Nuclear research receives only one

award—for research into the hard X-ray fission of light nuclei. Commenting on this prize, the Chairman of the State Prize Committee, Minister of Higher and Secondary Specialised Education, Vyacheslav P. Elyutin, stated that "this work is of exceptional significance for the study of thermo-nuclear reactions and radioactive fallout". Nevertheless, in view of the repeated acclaim given in the Soviet media this last year to the fast-breeder power programme, it is a little surprising that there is no technology prize for nuclear engineering. The only other mention of nuclear physics is in the "textbook" section—for a two-volume university text on experimental nuclear physics.

This separate section for textbooks is a relatively recent innovation. In the Soviet Union, textbooks are selected by the educational authorities on a Union-wide basis, so that, on occasion, criticism of an unsatisfactory text will lead to a discussion in top-level party journals. The demand for scientific personnel is expanding with each Five Year Plan—the current Plan requires an "output" from the universities and institutes of higher education of 9.6 million new scientists, technologists and specialists in the period 1976 to 1980, so the provision of suitable textbooks is a matter of ever-increasing importance. This year two other university textbooks, on pedology and invertebrate palaeontology, are also honoured, as is a secondary school text on economic geography.

Vera Rich

Recombinant DNA

THE final Environmental Impact Statement (EIS) on the NIH guidelines for research involving recombinant DNA molecules was published last week, more than 16 months after the release of the guidelines themselves. The formulation of the final EIS has been fraught with delays and uncertainties. An initial draft which came in for some heavy criticism was published over a year ago along with an invitation for informed comment. Drafting an acceptable final version has taken much longer than expected, and the publication schedule has slipped several times. This delay may have legal consequences. Friends of the Earth has filed a suit against the NIH and NSE in an attempt to halt funding of recombinant DNA research, charging breach of the National Environment Protection Act on the grounds that the EIS was not published simultaneously with the guidelines. The NIH counter this by pointing out that delaying publication of the guidelines would have led to greater delays in applying adequate safeguards to the research.

As for the EIS itself, there are few surprises. The document and its extensive appendices go over all the familiar arguments concerning the potential and postulated environmental impacts of recombinant DNA research and comes to the unsurprising conclusion that "the level of risk that will result . . . is acceptably small." This conclusion and the arguments used to support it will undoubtedly supply ammunition to both sides of the continuing debate.

Sandy Grimwade



Above: the Meridian Room at the Dunsink Observatory of the Dublin Institute for Advanced Studies, where a major fire early last October severely damaged the library and laboratories. Loss of instruments has meant loss of valuable work. But Professor P. A. Wayman, Director of the Observatory, is hopeful that the observatory will be restored. He sees as first priority getting the smoke damaged computer back into action. Restoration of the eighteenth century building which housed the library is likely, though not definite yet.

The Observatory was founded in 1783 and became part of Dublin University. In 1947 it was taken over by the Dublin Institute. It has no telescopes of its own.

Chris Sherwell, News Editor of Nature for the past two years, has taken up a post on the foreign desk at the Financial Times. Robert Walgate, until recently with New Scientist, takes Mr Sherwell's place. And David Dickson, recently with the Times Higher Educational Supplement, becomes our Washington News Editor at the beginning of 1978.

Austrian science minister attacks science journalists in fraud case

THE Austrian Minister for science and research, Dr Hertha Firnberg, released a 32-page document last Wednesday (16 November) attacking the science journalists of Austrian and other newspapers for poor judgment, and for unwittingly assisting in the execution of a crime. The attack, whether justified or not, threatens to disrupt the relationship—already fragile—between Austrian journalists, the ministry and the scientific community.

The government statement singled out two journalists for particularly harsh treatment: Dr Eleonore Thun-Hohenstein of the conservative daily *Wochenpresse* and Paul Uccusic of the middle-of-the-road *Kurier*. Their error? Publicising the work of a little-known freelance microbiologist, Herbert Schaden, and his assistant Hertha Celta. Schaden and Celta claimed that they had developed, or would soon develop, a strain of microbe which would digest plastic. No such strain emerged after four years' work.

Schaden was detained in Austria on 11 January, and has been in jail since, awaiting trial on charges of fraudulently obtaining funds for research (amounting to some £½ million) from a Dutch packaging firm and other sources. Celta was imprisoned with him but released on bail after two months—to feed Schaden's microbes. The microbes and Celta languish yet in Schaden's private laboratory at 99 Panzingerstrasse, a fashionable turn-of-the-century Viennese villa.

Dr Firnberg's document implies that Schaden is a crook—albeit a very clever one—and attacks the journalists for helping him along the path to fraud. Some 750 articles describing Schaden's work appeared in the press between 1 February 1973 and 5 June 1974, the document states, mostly in Austria, but also in Holland and in Italy.

Schaden first came to the journalists' attention on 1 February 1973, when he sent a thick bunch of documents to the Austrian Bundesministerium für Wissenschaft und Forschung, with copies to the Austrian newspapers. The documents gave credentials, references, and evidence that Schaden and Celta had found the plastic-digesting bug. Years before, a stack of plastic containers had been found riddled with microscopic holes; Schaden had deduced the presence of bacterial action and set about isolating the active agents.

On 21 February 1973 Thun-Hohenstein's first article, a positive one, appeared in *Wochenpresse*. A Dutch journalist noticed the story and

relayed it to her paper, the *Haarlems Dagblad*, a small but important provincial Dutch paper.

Oscar van Leer, a Dutch packaging magnate, came to hear of the story and was interested. Uccusic published his version of Schaden's work, also positive, in *Kurier* on 4 March. By 30 May Thun-Hohenstein had heard both that the Austrian ministry had dismissed Schaden ambiguously with a "don't ring us, we'll ring you", and that van Leer's firm (Royal Packaging Industries van Leer B. V.) was investing in Schaden's bugs. van Leer had sent a team of plastics experts—but not microbiologists—to interview Schaden and concluded he was genuine. Thun published a sarcastic article ridiculing the ministry's delay and poor judgment. On 9 June Uccusic, not a man to mince words, weighed in with his own criticism. He printed a story related by Thun: that the chief civil servant responsible for advising on the allocation of research funds in Austria—Dr Wilhelm Grimborg—had threatened Thun with a court action for saying the ministry's rejection of Schaden took five lines rather than the five and a half. Uccusic made great play of that. Later a letter from Grimborg was printed in *Kurier* rebutting Uccusic's accusations; but Uccusic tacked a stinging comment on the end.

Grimborg's relationship with Thun-Hohenstein and Uccusic was therefore very sour; the more so because the ministry's popularity with scientists was at a low ebb. Austria was—and is still—enjoying its first period of socialist government for a long while, and the ministry had embarked on a "demo-

cratisation" of the universities (for which it is responsible), introducing students and junior professors on to decision-making bodies, and dismantling the previous patriarchal and authoritarian structures. Criticisms of competence were particularly unwelcome at that time.

But now has come the retribution. van Leer invested a large sum in Schaden, only to find his work and his claim valueless. van Leer's lawyers believe they have discovered that some 40 pages of Schaden's original documentation are careful re-writes of earlier German papers, changing names and places. So the ministry appears to be vindicated (though Schaden and Celta's case has yet to come to court, and some feel that due process of law has been disturbed by the ministry's attack). Dr Firnberg, the minister, threatens wide distribution of the report, to all politicians and universities in Austria, an act which will undoubtedly severely weaken the position of the science journalists within their own newspapers and in Austria itself.

What of the journalists? Dr Thun-Hohenstein dismisses the Firnberg attack and indeed still believes in Schaden; Herr Uccusic thinks Schaden "not the person we took him for". He feels however that "neither I nor Dr Thun-Hohenstein nor anyone has any proof" that Schaden is wrong. And many journalists are saying that as it has taken experts since 11 January this year to develop a watertight case against Schaden—and may take longer—how can they have been expected to judge? The development was potentially of great public interest so had the public not the right to know?

Uccusic also takes a more philosophical view. "Error is always possible" he says; "even politicians and scientists can make errors". Uccusic believes scientific officialdom to be too hidebound—which it sometimes is—and he takes an interest in fringe science. But the belief can lead one into dangerous territory. Uccusic recognises the danger but adds "We should be open-minded and ready for new developments. In this special case no law of science forbids the development of organisms that might digest plastic. And Schaden was employing a recognised technique: to isolate a useful strain and strengthen it by multiplication and selection".

The final criticism must be laid at the door of the ministry itself. The ministry accuses the journalists of superficiality and lack of thoroughness. But the document they have produced in their haste to attack the journalists suffers from the same fault. It is too polemical; and its puts only one side of a difficult case.

Robert Walgate



Herbert Schaden and Hertha Celta, accused of obtaining £½ million by false pretences

EEC taps a megawatt of Italian sunshine

THE EEC Commission and a consortium of European companies signed a construction contract in Brussels last week for a £3 million electricity generating plant operated by the sun. The plant will probably be built in the south of Italy, and should be producing a megawatt of electricity by 1981—enough power for between 100 and 200 households.

Two hundred and fifty mirrors, covering a hectare of ground, will be used to reflect and concentrate the sun's heat on to a boiler, on top of a 50-metre-high tower, to produce steam to drive conventional turbines. The heat receiver will be about four metres square. Each square metre of mirror will generate 1 kW. The plant will have a thermal power of about 5 MW indicating a thermal efficiency of 20%. It will operate with steam at 510 °C and

64 atmospheres. The mirrors will be computer controlled to follow the sun and the whole structure will be solid enough to withstand winds of up to 130 km h⁻¹.

This will not be the first such solar electricity generating plant. France, which is taking part in the project, already has its own 2 MW plant near Marseilles and Macdonnel Douglas in the USA is planning to build a 10 MW plant in California at an estimated cost of \$31 million. But because of spending cuts that is likely to be delayed.

All these plants, however, will be expensive, experimental projects, operating on an uneconomically small scale. The maximum possible output for a solar plant is thought to be about 100 MW. Beyond that it becomes physically impossible to get more mirrors close enough to the boilers to

have any effect. Major modern electricity generating stations can produce 1,000 MW, and Mr Albert Strub, head of the EEC Commission's non-nuclear energy department, estimates that "if 5% of Europe's energy comes from the sun by the year 2000, we will be doing well".

The aim of the Italian plant is to teach the companies concerned the technical problems involved. The firms concerned are: the Italian national electricity company, ENEL; Ansaldo of Italy; Cethel (grouping the French companies Renault, St Gobain and Heurtey); and Messerschmitt-Bolkow-Blohm of West Germany. General Technology Systems of London will advise the Commission on the management of the project. Half the cost will be met from the Commission's research budget, the remainder coming from the participating countries.

Brian Donaghy

THE Royal Society of Tropical Medicine and Hygiene is holding a symposium this year, from 23 to 25 November, to celebrate the first centenary of medical entomology. It is just one hundred years since Patrick Manson showed that *Culex fatigans*, the common house mosquito of the tropics, was the intermediate host of a nematode worm which caused the disease filariasis in man. That was the first time an insect had been so involved, and although other scientists were engaged in parallel investigations, and there have been unpleasant confrontations between those supporting other claimants to priority, I have no doubt that Manson fully deserves this accolade.

The novelty of his discovery is manifest when we examine the records of his time. The more traditional members of the medical profession called him a lunatic, and intended to ridicule him when they gave him the nickname 'Mosquito Manson'. However, he inspired a few of his colleagues, particularly Ronald Ross. To the general public Ross, the discoverer of the transmission of malaria by the mosquito, was the more popular figure, but Ross himself always acknowledged his debt to Manson.

I do not think that it is any exaggeration to say that Manson's work has had a more profound effect on the human race than any other discovery which can be attributed to an identifiable individual. We cannot name the benefactors who discovered the wheel, the use of fire and the possibility of growing arable crops: their contributions may have been more revolutionary to man's way of

life, but the results are of the same order of magnitude. But for the work of Manson and those he inspired, the populations of many tropical countries might today be less than half their present level.

Mosquito Manson



KENNETH MELLANBY

A hundred years ago the majority of the world's population lived in areas where diseases which we now know to be transmitted by insects (and in no other way) dominated the lives, and the deaths, of the inhabitants, yet no one had the faintest idea how the diseases might be controlled. Even today malaria, typhus, sleeping sickness, filariasis, dengue and many more diseases are still important, claiming many victims and affecting all types of development, but we know how they could be reduced or even

eradicated. Though patchy, some degree of control has been widely implemented. The resulting rise in population may not be an unmixed blessing in a world of shrinking resources, but to many individuals, particularly in tropical countries, the improvement has been dramatic.

Although a hundred years ago no one appeared to believe that insects could carry diseases, it is tempting to search for earlier suggestions of such a mechanism. It is true that authoritative books on the English fens produced in the nineteenth century which devote much space to "the ague", which was generally malaria, give no inkling of any idea other than that the disease was caused by the "bad air" from the marshes. They discuss the importance or otherwise of water pollution by organic matter, which can now be interpreted as affecting the survival of mosquito larvae, but there is no mention of mosquito bites as a danger to health. However, I recently found a contemporary account giving an inventory of the contents of the Cistercian Abbey at Sawtry, only a couple of miles from where I write in Huntingdonshire, at the time of the dissolution of the monasteries in the sixteenth century. The Abbot's chamber had a feather bed and white curtains, but the guest room, reserved for important visitors, was hung with tapestry and a bedstead with "nets for knats". This may have been to protect the guests from infection with malaria, but I fear that it was just for his comfort and to keep out the insects which can still be troublesome on warm summer nights in fenland. But it may have kept him healthy all the same.

correspondence

Farm energy

SIR,—I am not sure what your correspondent Michael Knee is trying to say in his letter (13 October, page 556). He accuses me of "blind prejudice" for stating a few simple facts about the use of energy by farmers. If he reads my contributions in the issues of 25 August (Frugal farming) and of 8 September (Food for energy), he will find that he has misinterpreted my views.

I would be the last to deny that an enormous amount of energy is wasted by some modern farmers, but this is almost entirely because of the inefficiency of intensively-kept livestock, which waste over 90% of their food, and which is the main reason why Britain is not already largely self-sufficient. This is mentioned in my contribution of 25 August, and explained in detail in my book *Can Britain Feed Itself*. But this is irrelevant to the present argument.

My contention is that arable farming is efficient. It has been shown by everyone who has investigated the situation, including Gerald Leach in his *Energy and food production* (mentioned by Michael Knee) that arable farming, although using much energy to propel tractors and manufacture fertilisers, produces crops with at least three times the energy value of that used in growing them. Recent developments, such as direct drilling, will reduce energy expenditure without reducing the yields, and so further improve the equation. In fact intensive arable farming is probably the most efficient way we can trap solar energy in an easily useable form.

Like Michael Knee, I would like to see more people working on our farms, but men, like horses, would probably use more energy in a twelve-month than the present tractors and combine harvesters, if we were to grow and harvest comparable crops.

Yours faithfully,

KENNETH MELLANBY

Huntingdon, UK

Desert rainfall

SIR,—The statement by Glantz and Katz (19 May, page 192) with regard to measuring rainfall at two locations, Gao and Niamey, in the Sahel "that at least in some sense, the mean is too large and not at all indicative of how much rain commonly falls" is overly

strong. In no sense is the mean too large; the mean is simply the mean—a well defined and much used statistic. Furthermore it is quite indicative of how much rain commonly falls in most regions—even Gao and Niamey. Glantz and Katz also note that positive "skewness is characteristic of the rainfall distribution not only in the Sahel but, more generally, in arid and semi-arid regions. In particular the degree of skewness is greater the drier the climate". It has been our experience, however, that positive skewness is a characteristic of rainfall distribution for almost all localities and that, at most, there seems to be only a slight tendency for the degree of skewness to increase as the average amount of precipitation decreases.

We selected a random sample of fifteen reporting localities from the original *Climatic Summary of the United States* which contained data from the establishment of reporting stations up to 1930. The only restrictive requirement imposed in the selection process was that there had to be at least twenty consecutive years of records for a given locality. We used the mean, standard deviation, and coefficient of skewness for the distribution of precipitation at each location to examine the hypothesis that the more arid the climate the more skewed the distribution rainfall. If the hypothesis is valid, there should be a strong negative correlation between mean rainfall and the skewness coefficient. We found that there did seem to be more skewness for regions with little rainfall in the month of July than for a region of average rainfall, but positive skewness also seemed to be greater for regions with greater than normal rainfall. The correlation coefficient of $r = -0.08$ tends to confirm the lack of a simple systematic negative relationship between mean rainfall and skewness.

It also should be noted that positive skewness seems to be the normal condition for the distribution of rainfall for almost all localities and not just for the localities in arid regions. The meaning of this should be evident. If it is not true that the mean alone is suitable for describing the distribution of rainfall in arid regions, then neither is it alone a suitable measurement for describing the distribution in any region. This, however, is not a surprising result; few distributions can be de-

scribed by one parameter. As Glantz and Katz state, "no single number can adequately describe the climate regime of an arid or semiarid region". We would add that the statement holds for all regions and not merely for arid or semiarid regions.

We also would like to add a note of disagreement with the statement that, "recent weather tends to influence perceptions more heavily than earlier weather and wet spells more heavily than dry ones". While the statement about recent weather versus earlier weather seems totally acceptable, the latter part of the statement would seem to need at least some documentation. What evidence exists to substantiate this statement? If nothing else, it does seem evident that the drought of the past few years in the Sahel will influence perceptions and behaviour quite as much as the earlier wet spells.

Yours faithfully,

J. LARRY DEATON

United States Department of
Agriculture,
Washington DC

Using varied talents

SIR,—I was unpleasantly surprised at your negative editorial comment on Sir Andrew Huxley's Presidential Address (8 September, page 95). Accepting the idea that inherited differences exist does not automatically condemn the apparently less-endowed people to a subhuman status. Those with no apparent talent are not necessarily inferior; they simply can't find their true calling in our culture. A century ago born atomic physicists and computer programmers may have spent their lives on the skid row as impractical silly dreamers; today they are productive, well-regarded members of our society.

As civilisation advances, more and more people will find the calling that uses their innate talents; fewer and fewer will spend their lives as frustrated misfits. All we have to do is to maintain progress in our civilisation, so more and more varied talents can be used. However, to maintain this progress we must have the courage to face the apparently impalatable as well as the palatable results.

Yours faithfully,

ANDREJS BAIDINS

Du Pont Company,
Wilmington, Delaware,
USA

news and views

DNA insertions and gene structure

from Bob Williamson

ONCE again we are surprised, this time by the structure of genes coding for proteins in higher animals. The primary sequence of the protein is determined by messenger RNA (mRNA); this in turn is synthesised on one of the strands of genomic DNA, probably first as a precursor of high molecular weight (HnRNA). Several animal cell mRNAs have been sequenced in whole or part, revealing no peculiarities so far: the coding sequence is preceded by a capped, non-translated region, starts with an initiator triplet, and is read in phase to the 3'-terminator triplet, which is followed by a non-translated 3' region and last, a polyadenylic acid sequence. The sequence of the coding region of the mRNA exactly specifies the amino acid sequence of the protein when read in register.

It was assumed that the mRNA was copied directly from a gene sequence, and data from prokaryotes certainly supported this view—a frame-shift mutation in which a base was added or deleted caused the predicted change in protein sequence, and even, in an operon, in proteins coded adjacent to the 3' side of the genome. In animal cells, however, this now turns out not to be the case, at least for ovalbumin and β -globin genes.

These genes are particularly suitable for study as they are present only in single copies, and tissues which make each protein in large amounts can be compared with those that make little or none. The specific mRNAs are available, and the construction of recombinant DNA plasmids has removed the problems of purity and amount that particularly apply to single-copy genes because they are present in the sequences in nuclear ('genomic') DNA only as one part in several million.

No sooner was an appropriate recombinant clone of DNA available (courtesy of T. Maniatis) than Flavell and Jeffreys started the analysis of the genes for rabbit β globin. They 'mapped' the gene—that is they found the size of the DNA fragments containing the gene when the DNA was digested with various restriction enzymes by techniques developed by Southern (*J. molec. Biol.* **98**, 503; 1975). They looked for the nearest repeated DNA sequences, the nearest short poly(dA) runs, and so on. It all made sense for a while—the gene for β globin (there are actually two related genes, which seem to correspond to the β - and δ -globin genes in humans) at first behaved correctly, and a picture of the gene itself, the surrounding bits of DNA and the minimum distances to other genes became partially clear (Jeffreys & Flavell, *Cell*, **12**, 429–441; 1977).

Then more enzymes were used. When a variety of enzymes is used, in a way analogous to that in which proteases are used to determine protein sequences, they have to 'add up'—if sequence AC is cut at point B, then AB plus BC has to equal AC. This is so 'self-evident' that Jeffreys and Flavell deserve the greatest praise for not dismissing their data out of hand when they showed that 2+2 does not always equal our preconception of 4.

In the β -globin mRNA, the sequence of bases starts at one end with the triplet specifying the N-terminal amino acid of the protein and moves steadily and without interruption to the other end, where the protein terminates. But in the gene the sequence of bases coding for β globin is not continuous. In particular one fragment of cloned complementary DNA (obtained by reverse transcription of the mRNA) which is 333 bases long, derived entirely from the coding sequence and bounded by two *Hae*III restriction sites, hybridises to a *Hae*III restriction fragment of genomic DNA that is more than twice

as big. Clearly the restriction fragment contains an additional insert (some 700 base pairs long). The only conclusion must be that there is a bit of DNA, as big as the structural gene for β globin itself, in the middle of the gene, which does not appear in the final mRNA (Jeffreys & Flavell *Cell*, in the press).

At the same time as Flavell and Jeffreys were obtaining this startling result, Leder *et al.* were recording a related and equally inexplicable finding (*Cold Spring Harb. Symp. quant. Biol.*, in the press). Leder's group had first to prepare a safe vector to clone mammalian genes, and chose bacteriophage λ . They introduced a number of mutations and deletions so that only recombinants could be 'packaged' and so reproduce, and also several mutations which rendered the phage non-viable outside controlled laboratory conditions. They then purified the gene for mouse β globin several hundred times using a combination of reverse phase chromatography (which separates on base composition and secondary structure) and Agarose gel electrophoresis, and obtained two fragments approximately 15,000 and 7,000 base pairs (15 kb and 7 kb) long. Each of these, while much purified over the original DNA, still only contain one globin gene-containing sequence in every thousand or so total sequences.

The mixture was then recombined with the disabled λ vector. Recombinants containing the mouse β -globin gene were selected. The length of the cloned mouse DNA was ten times the size of the coding sequence, some 7 kb long, and contained the entire β -globin gene. However, once again, it was not all in a single piece, for when it was hybridised to mouse globin mRNA and examined with the electron microscope, a remarkable extra loop in the middle of the DNA was found. Restriction enzyme analysis confirmed the existence of the insert. This insert is 450 base pairs long—it is not yet clear whether it is in the same position as

for the rabbit insert described by Flavell and Jeffreys, nor whether its sequence is conserved.

The chicken ovalbumin gene has been studied almost as extensively as the globin gene, largely because it is hoped to learn from it more about gene action in response to hormone stimulation. Because it is present only in a single copy, fragments of genomic DNA prepared using a specific restriction endonuclease which does not break the gene should once again give a single DNA fragment containing the gene—if the sequence is present as a single unit in the DNA as in the mRNA. However, two groups have just shown that as for globin, several fragments rather than one are often found (Breathnach, Mandel & Chambon, this issue of *Nature*, page 314; Doel, Houghton, Cook & Carey, *Nucleic Acids Res.* **4**, 3701–3713; 1977).

Chambon and his colleagues first prepared a recombinant plasmid between complementary DNA (cDNA) prepared from chick oviduct ovalbumin mRNA and pCRI. This enabled the Strasbourg group to determine which restriction endonucleases did not cleave the coding regions of the gene.

A slightly different approach was used by Carey's group at Searle. They first prepared very pure mRNA and isolated full length cDNA, and then synthesised a second complementary strand with reverse transcriptase. Full length double strand cDNA was isolated by Agarose gel electrophoresis. This too could be checked for restriction enzyme sites in the coding gene region. (Carey *et al.* also checked their results against the plasmid prepared by the Strasbourg group, a good example of international co-operation in the use of recombinant DNA.)

Neither of the restriction enzymes *EcoRI* nor *HindIII* cleaved the structural gene sequence in Chambon's plasmid or Carey's double-stranded cDNA. These then were the enzymes used to fragment genomic DNA from chick tissues.

Both groups, using their different techniques, found three DNA sequences which hybridised to either cDNA or nick-translated plasmid DNA after restriction of genomic DNA with *HindIII*. Using *EcoRI*, Chambon *et al.* found convincing hybridisation in four positions. Thus instead of a single band expected, several gene-containing regions separated by long inserts, or spacers, are found. Moreover, this is true for all tissues studied, whether making ovalbumin or not.

In the case of mouse and rabbit β globin, the entire coding gene was isolated, with the insert in the middle. The structural gene regions were therefore at least closely linked, if separated by the unexplained sequence. In the case of ovalbumin, however, no frag-

ment containing both coding sequences has been found and one of the inserts is at least 7,000 base pairs long. It is even theoretically possible that different regions of the ovalbumin structural gene are on different chromosomes, and not linked at all.

How can this remarkable result help in the understanding of transcription? Very little, until several questions have been answered. As Chambon *et al.* point out today, there are three apparent possibilities. The 'primary' transcript (the HnRNA containing the mRNA sequences) may include the inserts and all the mRNA coding regions, with the inserts being specifically nicked out and the mRNA ligated to its final form in processing. This is quite possible because, as has been stated here (*News and Views* **269**, 648; 1977), the size of the globin 'precursor' most easily identified for β mRNA is 15S, just the size expected if the insert is still present flanked by 3' and 5'-coding regions. However, such a correlation is at best indicative. The concept of a transcribed insert is also, however, supported by the puzzling fact that both the poly (A) sequence at the 3' terminus of the mRNA and the methylated 'cap' at the 5' end seem to be conserved from the HnRNA to the final polysomal mRNA. The explanation of this is clear if the processing occurs not from one end or the other, but out of the middle.

If the inserts are transcribed, there must be a very accurate enzymatic mechanism for cutting them out. Such excision enzymes are known (Robertson & Dickson, *Brookhaven Symp. Biol.* **26**, 240–266; 1975), but have never been demonstrated in animal cells, although ribosomal RNA processing certainly involves similar accurate cleavages. The inserts might be expected to contain extensive self-complementary double-strand regions, such as palindromes, but they do not self-reassociate (at least for rabbit β -globin genes), indicating that if such sequences are present, they are quite short. Also, the inserts may be repeated a few times throughout the genome (Flavell and Jeffreys feel 10–15 times is the highest reiteration frequency their data might support), but are certainly not common to many structural genes. Perhaps the excision specificity is mediated by nuclear protein which, in interacting with HnRNA, would both force certain RNA conformations and also act as additional markers for specific enzymes—or even have enzymatic activity themselves.

The second possibility is that looping out occurs at the DNA level, and the inserts are never transcribed at all. The main attraction of this idea is that, while we know little about DNA structure in active chromatin, most pos-

tulates are still possible. (It also would help to explain the difficulty in obtaining convincing and reproducible transcription from chromatin because additional factors would be required to control DNA conformation for active genes.)

Finally, it is possible that the portions of the mRNA are synthesised separately and then ligated together during processing. In this case the 'inserts' might or might not be transcribed, but the various bits of the coding sequence would not be co-transcribed in a single long piece of RNA. Interestingly, the gene for which the best evidence exists for the absence of HnRNA is precisely that reported here—McKnight and Schimke (*Proc. natn. Acad. Sci. U.S.A.* **71**, 4327–4331; 1974) demonstrated that it is very difficult to show an ovalbumin HnRNA, using techniques which seem to work for globin.

No doubt we shall soon know whether the 'insert' is transcribed, and perhaps have definitive data on its multiplicity in the genome and whether it includes short palindromic signals which might indicate looping-out sites. Such experiments will be most interesting for globin, where chromatographic techniques allow the isolation of precursor RNAs (*News and Views* **269**, 9; 1977).

It is very interesting to compare these data for globin and ovalbumin genes with that reported last month in these columns (*News and Views* **269**, 648; 1977) for mouse immunoglobulin genes. In the case of Ig, there are also inserts, in DNA from antibody-synthesising tissues and from other tissues, but the inserts appear different, and at least some somatic gene reassortment takes place. For globin and ovalbumin there is no apparent difference between tissues making the mRNA and those completely inactive for the protein in question. The inserts may allow novel somatic recombination events to occur, but do not make them obligatory. It is also unclear whether these sequences are related to the leader sequences which are transcribed from remote regions of the genome in adenovirus and SV40 (*News and Views* **268**, 101; 1977).

Let us briefly consider some of the wilder possible implications of this finding. If transcripts are ligated together during processing, there is no need that they be arranged in the 'correct' order—from 5' to 3'—in the genome. In fact, genes could exist with a single 5' end and several 3' ends, which are used interchangeably or in sequence during development. Maybe the β - and δ -globin genes or the two α -globin genes share some parts in common, only the 'different' bits being ligated in the right proportions. If transcribed, the inserts must be (how

easily we fall into such formulations, in spite of these findings) nicked out intact; the more reason to think they could have other regulatory functions. However, if not transcribed, surely as indicated above, the inserts are now the best candidates for marker sites for binding gene-specific non-histone proteins, and thus determining chromatin conformation during the control of transcription.

Those of us who are interested in safety with recombinant DNA sequences are left with the following thought. Until now, it has been assumed that cDNA derived from an mRNA is less likely to be expressed

(and therefore, in some contexts, might be regarded as 'safer') than a total gene sequence, including the flanking regions. Now with the discovery of gene inserts it is quite likely that the controls involved in transcribing and/or processing several different bits of DNA into highly complex precursors of RNA, followed by excision and ligation steps, will be beyond the capabilities of the bacterial hosts used for recombinant DNA vectors. Hence the genomic DNA recombinants may prove to be just as safe as cDNA recombinants unless much work goes into inserting processing systems as well. □

Molecular hydrogen in the Sun

from A. H. Cook

A RELATIVELY large proportion of matter in galaxies lies in the space between stars and until fairly recently appeared to be in the form of free atoms and solid grains. Atoms were detected in two ways—by the microwave radiation from the hyperfine transition in the ground state of hydrogen at 1,421 MHz and by absorption of starlight by atoms showing up the spectra of the stars in the visible region. Solid particles scatter, polarise and absorb visible light from stars. There was also evidence for the existence of a few molecules, CH and CN in particular, absorption lines of which lay in the visible. It was also well known that some molecules, such as C_2 , TiO, and CN occurred in cool stars. By and large however, limitations on the sensitivity of radio measurements and the restriction of optical spectroscopy to the visible meant that direct evidence of the existence of molecules could not be obtained. At the same time, dynamical considerations indicated that there was more mass in the Galaxy than could be accounted for by the observed hydrogen and other atoms.

So far as interstellar material goes, there are now extensive observations of molecules. Radio observations show absorption or emission by almost 50 molecules, a few of which radiate by maser action. Ultraviolet spectra of stars especially those obtained by the Copernicus satellite (Morton *Phil. Trans. R. Soc. A* **279**, 299; 1975) show

absorption by the H_2 molecule. The abundance of interstellar H_2 molecules can also be estimated from the radio observations of carbon monoxide. It is now plain that molecules are important in two ways. Many different molecules are observed in the clouds of gas that seem to be condensing into stars and the total mass of hydrogen as molecules and atoms is much greater than that as atoms alone.

Are molecules as important in stars? Clearly they are significant in the outer layers of cool stars. Now, in this issue of *Nature* (page 326) Jordan *et al.* describe their identification of H_2 in the ultraviolet spectra of sunspots. The lines are in emission between 120 and 160 nm, and only a few of the possible lines of the H_2 spectrum in that range are observed. Jordan and others show that the observed lines are just those to be expected if a particular level is preferentially populated by the absorption of the Lyman- α radiation of atomic hydrogen, which is of course very intense in the solar atmosphere. Selective fluorescence, as this process is, is of course familiar in laboratory studies. It is also known in astrophysics.

The spectra of molecules in comets are fluorescent spectra excited by sunlight, but again, not all possible lines occur. As was shown by Swings, the reason is that some energy levels are not excited because the frequencies of the transitions by which they would be excited from the ground state coincide with strong Fraunhofer lines in the solar spectrum. That is the reverse of the present case, where the upper level in the H_2 spectrum is populated by absorption of a single line, Lyman- α ,

which is much the strongest emission line in the solar spectrum. The coincidence between the Lyman- α frequency and that of the nearby line in the H_2 spectrum is not exact but the Lyman- α line is so broad that there is overlap.

It is natural to ask whether there are other molecular fluorescence spectra waiting to be found in astrophysics. A general argument suggests the chance is high—molecular energies and transitions are numerous because of the many rotational and vibrational states, so that the chance of a coincidence with a transition of an atom or another molecule is much higher than for atomic spectra. It is, for example, not out of the question that a specific fluorescence process might be responsible for inverting populations of levels in the ground state of a molecule from which microwave stimulated emission occurs (Cook *Nature* **210**, 611; 1966); indeed the process suggested by Litvak (*Astrophys. J.* **156**, 471; 1969) for infrared pumping of the hydroxyl maser is of this type, although it relies on the hydroxyl itself to filter continuous radiation and produce a distribution of intensity that gives rise to selective fluorescence. Furthermore, as Jordan and others themselves mention the selective absorption of Lyman- α by interstellar molecules may be important in the formation of interstellar spectra of molecules (Osterbrock *Astrophys. J.* **136**, 359; 1962; Black & Delgarno, *Bull. Am. Astron. Soc.* **6**, 444; 1974).

The other question which is raised by the analysis of Jordan and others is the part played by molecules in stellar atmospheres. If molecules are at all abundant they will affect the absorption of radiation in an atmosphere in two ways. In the first place, there will be continuous absorption leading to dissociation and the frequencies for dissociation are in general much less than those for ionisation of atoms or molecules. Second, molecular lines often lie so close together because of the many rotational and vibrational states that exist, that in some parts of the spectra absorption in discrete lines will come close to continuous absorption. Molecules will of course be dissociated if the temperature of the atmosphere is high but many molecules can exist at temperatures up to 3,000 K. It has indeed been suggested that hydroxyl may contribute to the opacity of the solar atmosphere, and it is almost certain that H_2 is present in cool stars where molecules such as C_2 and CN are already observed.

The study of Jordan and others and the ideas which stem from it call attention to the importance of molecular spectra in the neighbourhood of Lyman- α . Because the energies lie above the dissociation energy of most

molecules, predissociation phenomena may be observed in the spectra. The advent of synchrotron sources for ultraviolet spectroscopy makes it easier to investigate molecular spectra in the far ultraviolet, and prompted by the increasing astrophysical interest in molecules, it may be that there is an

area of physics here ripe for development. Thus the study of Jordan and others calls attention to these aspects of molecular astrophysics—the role of molecules in stellar atmospheres, the formation of molecular lines in interstellar formation and the expanding interest in laboratory investigations. □

How should earthquakes be quantified?

from a Correspondent

At the meeting of the International Association of Seismology and Physics of the Earth's Interior (IASPEI) held in Durham on 9-19 August, 1977, a special workshop on the subject of Quantification of Earthquakes was organised by Professor K. Aki (Massachusetts Institute of Technology).

MUCH of the discussion in this workshop was concentrated on the concept of seismological magnitude, the oldest and most widely used measure of the size of earthquakes. Although magnitude has been an extremely useful phenomenological tool, recent advances in seismic source theory have convinced many seismologists that it is, if not obsolete, at least in need of revision and redefinition. There seems to be no agreement about how magnitude can be related to the physics of earthquakes or even about what it is that magnitude is intended to measure.

Also at Durham, the Subcommittee on Magnitude of the Commission on Practice met and wrestled with the problem of organising and simplifying the profusion of differing magnitude scales in current use. This task is almost impossibly difficult, due to variations in instrument design, wave types used, seismogram-reading practices, frequency and distance correction methods, and nomenclature. The situation has become so complex that even to catalogue current practices has required a major effort and has resulted in a document of impressive size. If the Subcommittee succeeds even partially in reducing the confusion and in standardising practices, it will have performed a valuable service.

Seismological magnitude was originally developed by Gutenberg and Richter in the 1930s as a heuristic tool for studying Southern California earthquakes. Before then, statistical studies of earthquake occurrence had made no distinction between large and small events, and often produced very misleading results. Magnitude was defined as proportional to the logarithm of the maximum excursion produced on a Wood-Anderson torsion seismograph

situated 100 kilometres from the epicentre, with magnitude 0 set at a level near the smallest detectable earthquakes. (Magnitudes down to about -3 are nowadays reported in micro-earthquake studies.) In this definition, Gutenberg and Richter were influenced considerably by the stellar magnitude scale in astronomy, from which seismological magnitude differs in two respects. First, and most trivially, the scales run in opposite directions: bright stars have small magnitude numbers, but great earthquakes have large ones. More importantly, seismological magnitude is intended to measure only the size of earthquakes, while astronomical magnitudes are of two varieties: absolute, which measures the total radiation from stars, and visual, which measures their brightness as seen (usually) from the Earth. The closest thing seismology has to visual magnitude is intensity, which is a measure of the strength of ground shaking in the epicentral region. Seismological intensity, however, is not based on any kind of measurement, but rather on human judgement applied to the social and engineering effects produced. It does not lend itself to quantitative interpretation, and in fact is usually expressed in Roman numerals to discourage any such attempt. As strong-motion seismographs become more widely distributed, it will become possible to place intensity on a more objective foundation. A first step in this direction might be to establish an instrumental scale for intensities below the threshold of perceptibility, where seismographic data are plentiful.

Gutenberg and Richter themselves seem to have been surprised at the great success of the magnitude concept, much of which no doubt stems from the fact that the range of magnitudes is enormous, so that even an imprecise scale is still quite useful. Detectable earthquakes range over more than 11 orders of magnitude, while even a gross blunder such as doubling the observed signal amplitude adds only 0.3 to the magnitude. By now however, our understanding of earthquake mechanisms has advanced to the point where

several physical attributes can be determined, at least approximately, from the radiated waves (such as source dimensions, displacement dislocation and stress drop). The relation between any of these quantities and magnitude is purely statistical, and subject to a great deal of scatter. (There is a case in California of a magnitude 3 earthquake with a fault length usually typical of magnitude 6 events.) It has become clear that several numbers are needed to describe an earthquake's size, and that magnitude is not one of them.

One of the most serious deficiencies of magnitude as it is now defined is the phenomenon of saturation. To determine the size of an earthquake, it is necessary to use waves longer than the spatial dimensions of the source. Otherwise, one measures only the radiation from a small part of the total source region at each instant. As a result the magnitude measured in any given instrumental pass-band has a practical upper limit. For 'short-period' (~1 s) waves this limit is around magnitude 7, while at 20 s it is about 8.5. If magnitude is based on the longest observable waves, then values range at least as high as 10 (for the 1960 Chilean earthquake). One solution to the saturation problem is to use waves longer than the source dimensions (how long this may be will vary from event to event). The quantity so measured is referred to as 'seismic moment', and its use has spread rapidly since its introduction some 10 years ago. It has the additional advantage of being subject to a simple physical interpretation: for a shear fault, for example, it is the product of the rigidity modulus, the fault area, and the mean dislocation across the fault. Seismic moment suffers, however, from being based only on low-frequency motions. In many contexts, such as that of hazard evaluation, high-frequency motions are more important.

Another related limitation of magnitude is the fact that it is based on instantaneous measurements. Signals of the same amplitude, whether they last for 1 cycle or 100, always lead to equal magnitudes, though they certainly do not correspond to similar source processes or hazards. Recently, local earthquake magnitude scales based on signal duration, or 'coda length', have been developed and are used widely in California, but they have not yet been given a sound theoretical foundation; the coda length observed is primarily a wave-propagation (scattering), rather than a source effect and may be subject to large variations with seismic region and focal depth. Coda-length magnitude has so far been used more to avoid problems of instrument saturation than to study the actual duration of earthquakes.

Considerable attention in the workshop was given to a generalisation of seismic moment, the 'seismic moment tensor.' This representation of seismic sources, which was suggested by Gilbert in 1970 (*Geophys. J. R. astr. Soc.* **22**, 223) and has recently been justified rigorously by Backus and Mulcahy (*Geophys. J. R. astr. Soc.* **46**, 341; and **47**, 301), promises to narrow the gap between observational seismology and the theory of earthquakes. The problem of understanding the earthquake process is one of the most difficult in mathematical physics, involving the elastic and anelastic behaviour, chemistry, and thermodynamics of rocks, and their interaction with the dynamic stress field. Seismologists are still far from solving this problem, and are likely to remain so for some time. When earthquakes are better understood, it may be possible to 'invert' seismic observations to derive event parameters, but meanwhile phenomenological descriptions will be needed, if only to catalogue observations we do not yet understand. This need explains much of the success of the magnitude concept, as well as the promise of the moment-tensor formulation. Backus and Mulcahy have shown that a unique reconstruction of the elastodynamic source process cannot, even in theory, be made from the information carried by the radiated waves; different physical sources can excite identical waves. What can be determined is the moment-tensor density, as a function of position and time, and any indigenous source can be described this way. To go further, unjustifiable assumptions must be made about the failure mechanism. The use of the moment-tensor representation, in fact, greatly simplifies the determination of source parameters, because the wave field depends linearly on the moment-tensor components, so the methods of linear inverse theory can be used. As a practical matter, this transformation of a non-linear inverse problem into a linear one enormously simplifies computations. The moment-tensor representation thus has the three advantages of being completely general, of being the most detailed source description which does not depend on a detailed model of the failure process, and of being computationally convenient.

It seems certain that moment tensors will prove to be powerful tools for advancing our understanding of the earthquake process. Agencies producing routine earthquake catalogues, such as the US Geological Survey and the International Seismological Centre, will probably soon begin to report moment tensors for large earthquakes, and these may someday replace magnitudes as the standard measure of earthquake size. In any event, though, seismo-

logical magnitude will be with us for some time, and its ultimate replacement or modification will occur because of scientific advances, not the reports and resolutions of official committees. □

New earthquake magnitude scale

from Peter J. Smith

ALTHOUGH more than a million earthquakes occur each year, most of the energy released comes from the very few very large ones. There is some point, therefore, in being able to determine the seismic energy from very large earthquakes with reasonable accuracy. Unfortunately, there are difficulties in doing so. It is usual to estimate seismic wave energy (E) from an earthquake's surface wave magnitude (M_s) using the Gutenberg-Richter equation $\log E = 1.5 M_s + 11.8$ or very similar empirical relationships derived by others (most notably Bath). The problem is, however, that for a great earthquake, which in this context means one with a rupture length of more than 100 km, the measured M_s is suspect.

The reason is that M_s is conventionally determined using surface waves with a period of about 20 s or by empirical conversion from a body wave magnitude derived from waves with even shorter periods. But when the rupture length of an earthquake exceeds the wavelength of the seismic waves used in the magnitude determination, the measured magnitude no longer reflects the entire rupture process; and there is then little correlation between M_s and rupture length. So the M_s scale is subject to 'saturation' at its upper end, the result of which is that for great earthquakes the measured magnitude and the energy estimated from it may be inaccurate.

In an attempt to circumvent this difficulty, which has been known for some time, Kanamori (*J. geophys. Res.* **82**, 2981; 1977) has now adopted a different approach to the calculation of the energy released by a great earthquake, an approach which leads incidentally to a more appropriate magnitude scale for such events. He thus begins not with magnitude but with seismic moment, a static source parameter defined as the product of the rigidity of the rock in the fault zone, the average slip over the fault surface and the area of the fault slip.

Peter J. Smith is a Reader in the Department of Earth Sciences at the Open University.

Seismic moment (M_0) is essentially a measure of the overall deformation at the earthquake source, but it can also be related to the elastic strain energy (W) released by the earthquake. It turns out that if all the built-up stress is removed by an earthquake, $W = W_0 \sim M_0/2 \times 10^4$. If the stress drop is not complete the expression for W becomes more complicated and the calculated W_0 gives only the minimum strain energy release. However, the situation is simplified if it be assumed, with Orowan (*Geol. Soc. Amer. Mem.* **79**, 323; 1960), that the stress remaining after the earthquake is equal to the frictional stress during faulting. In this case W_0 is not only the minimum estimate of W but is also equal to the seismic wave energy (E).

It is not clear whether or not the stress drop during an earthquake is complete; there is evidence to support either view. But there is also evidence from several earthquake studies to support the validity of Orowan's condition, which, if accepted, allows E to be determined from W_0 and hence from M_0 , irrespective of whether W_0 is the actual or the minimum strain energy release. As for M_0 itself, this quantity has been determined directly for many of the great earthquakes that have occurred since 1904, for others it may be estimated from the area of the fault plane or by other means, and for a few it may no longer be estimated at all. But whether M_0 is known, by whatever means, Kanamori has determined W_0 and hence E . By putting these E values into the Gutenberg-Richter equation he has then calculated magnitudes, denoted generally by M_w . In other words, M_w represents a new magnitude scale for great earthquakes, determined not directly from wave amplitude but indirectly from seismic moment.

The result is interesting. For the period 1904-76 the event with the highest M_w (9.5) turns out to have been the Chilean earthquake of 1960 ($M_s = 8.3$). This was followed by the Alaskan earthquake of 1964 ($M_w = 9.2$, $M_s = 8.4$), the Aleutian earthquake of 1957 ($M_w = 9.1$, $M_s = 8.3$) and the Kamchatka earthquake of 1952 ($M_w = 9.0$, $M_s = 8.3$). But M_w is not always greater than M_s . For the 1920 Kansu (China) shock, for example, an M_s of 8.5 becomes an M_w of 7.8. Moreover, for many earthquakes with a rupture length of about 100 km, M_s and M_w are in fairly good agreement, suggesting that for these and smaller events the M_s scale is entirely appropriate. This also means that the M_w scale connects smoothly to the M_s scale to which it becomes a natural continuation.

The M_w scale seems to be better for earthquakes because it removes the saturation inherent in the M_s scale.

But it also draws attention to some possibly significant relationships which either were not previously apparent or are different from those implied by the use of the M_s scale. For example, when W_0 is plotted against time for the period 1920–76 (for which the data are almost complete) it becomes clear that the annual average for the period 1920–65 (within which W_0 peaks) is more than an order of magnitude higher than that for the periods 1920–50 and 1965–76. This picture is quite different from that based on E calculated from M_s , which shows a steady decrease since the mid-1940s. The significance of this difference becomes clear with the discovery that the E -from- M_s curve is very similar in shape to the corresponding time curve of the number (N) of earthquakes with $M_s \geq 7.0$. Since the N curve is heavily biased in favour of moderate-to-large earthquakes, the proportion of events

with $M_s > 8.0$ being small (3.3%), the implication is that the E -from- M_s curve is also biased in favour of moderate-to-large earthquakes, indicating that for great earthquakes the seismic wave energy calculated from M_s is indeed underestimated.

This being the case, it becomes clear from a comparison of the W_0 and E -from- M_s curves that the number of, and the wave energy radiated by, moderate-to-large earthquakes decreased sharply just when the energy from great earthquakes increased sharply. Was this just a coincidence or was there a causal link there somewhere? The evidence suggests the latter. As Kanamori demonstrates, there is a remarkable correlation between the W_0 curve and the curve of Chandler Wobble amplitude, with W_0 moving in the same way as, but slightly behind, the amplitude.

One possible explanation for this correlation is that an increase in Chandler Wobble amplitude, brought about by external (for example, atmospheric) forces, triggers worldwide seismic activity and accelerates plate motions, eventually leading to great plate-decoupling earthquakes which so decrease intraplate and interplate stresses that moderate-to-large earthquake activity declines. On the other hand, the activity of great earthquakes could produce the Chandler Wobble, although this seems to be less likely. Then a third possibility is that the Chandler Wobble and great earthquakes have a common cause such as perturbations in the Earth's rotation. Whatever the explanation may be, the ancient controversy over the precise connection, if any, between earthquakes and Chandler Wobble has evidently been revived. □

Loess is essentially a silty soil, widely believed to have been deposited by the wind and is of particular interest as the soil on which some of the earliest agriculture flourished. The aims of the Loess Commission have been to produce a map showing the distribution of loess in Europe and to untangle some of the stratigraphic problems presented by the thick loess sequences of central Europe. It was formed as a subcommission at the 1961 INQUA meeting in Warsaw and upgraded to full commission status at the 1969 Paris meeting. Julius Fink (Austria) was its first president and he has steered the commission to the virtual completion of its initial tasks (see *Eiszeit. u. Gegenwart* 27, 220; 1976). At Birmingham, progress on the loess map was reported by G. Haase (DDR); the western sheet is virtually complete and proof prints should be available by the end of 1977. The eastern sheet has been delayed by the lack of a suitable topographic base but with the help of I. P. Gerasimov (USSR) this difficulty has now been overcome. Details of the map scales and the data presented are given in a paper by Haase, Ruske (DDR) and Fink which will appear soon in *Petermanns Geog. Mitt.*

Fink reported on field trips since the last congress (Christchurch 1973); visits have been made to Germany (BRD, 1974), France (1975) and the Soviet Union (Ukraine and Moldavia, 1976). Now that the European problems have been tackled the commission is looking east to the Asian loesses and the 1976 trip to Soviet territory represents the first tentative

New look for Loess Commission

from I. J. Smalley

Commission 4 of the International Union of Quaternary Research (INQUA) met at Birmingham University on 16 and 23 August, 1977. The scope of the Loess Commission is to be expanded, and practical aspects of loess investigation will be emphasised.

move in that direction. The commission will also turn to the study of more practical aspects of loess and give some emphasis to irrigation problems and engineering and economic topics. New INQUA bye-laws require sweeping changes in the commission membership and advantage will be taken of this to reflect new interests (both topical and regional). The new president is Marton Pecs (Hungary), Director of the Geographical Institute of the Academy of Sciences in Budapest; vice-president, B. Frenzel (BRD); secretary, O. Fränze (BRD); and a group representing Europe: J-P. Lautridou (France), J. Fink (Austria), J. Macoun (Czechoslovakia), A. E. Dodonov (USSR) and I. J. Smalley (UK). R. V. Ruhe (USA) will co-ordinate activities in North America and J. M. Bowler (Australia) will do the same for the Australasian and Pacific regions; there are eighteen corresponding members.

J. M. Bowler reported on the contacts between Australia and China. A group from the Australian National University has visited the classic loess

regions of north China and returned with samples and Chinese loess literature. Important papers will be translated and published in English. In particular the major works by Liu Tung-sheng and his collaborators will be made available. Chinese investigators have visited Australia to look at aeolian sediments in the central regions. There is no loess as such in Australia but the aeolian sediments have enough features in common to make comparative studies worthwhile. The contact at present is established at governmental level but the commission hopes to participate in studies of one of the world's most fascinating and still problematical (see *News and Views* 267, 484; 1977) loess deposits.

Pecs rounded the meeting off with a personal view of new aims for the commission. He stressed the need for research into practical problems connected with loess and proposed that in the next phase of the commission's activities there should be some emphasis on foundation problems and related engineering topics, on problems of landscape deterioration in loess areas, and on problems of forecasting and planning for land use. Traditional activities should not be neglected however and it might be possible to investigate certain loess stratotypes very thoroughly—perhaps on an international basis. He proposed that the section at Paks in Hungary might receive the same intense study as that organised by Fink at Krems in Austria.

The next loess discussion will probably be in Jerusalem at the Sedimentological Congress in July 1978. D. H. Yaalon (Israel) will prove to doubters that desert loess exists. □

I. J. Smalley is a Lecturer in the Department of Civil Engineering at the University of Leeds.

review article

Evolution of ageing

T. B. L. Kirkwood*

An evolutionary view of ageing suggests that mortality may be due to an energy-saving strategy of reduced error regulation in somatic cells. This supports Orgel's 'error catastrophe' hypothesis and offers a new basis for the study of normal and abnormal ageing syndromes and of apparently immortal transformed cell lines.

THE question of why we and most other higher animals age is not trivial since many organisms, notably higher plants, live and propagate indefinitely. Perhaps because we are so familiar with ageing it is dangerously easy to fall into circular argument such as that ageing is a mechanism for ridding a population of old and worn-out individuals who would otherwise compete for resources with younger and fitter ones¹. This makes sense if, for example, we assume individuals to have a defined reproductive lifespan, but is circular since the cessation of reproduction is itself a phenomenon of ageing. In this paper I show that by adopting an evolutionary point of view it is possible to relate Orgel's theory² that ageing is the result of a progressive breakdown in accuracy in protein synthesis to recent theoretical and experimental work on the senescence of diploid human fibroblasts, and I suggest a new reason why we may expect the error theory of ageing to be correct.

In discussing the evolution of error regulatory mechanisms it is appropriate to begin with brief consideration of the emergence of the earliest life-forms. The fate of the first replicating organic molecule would have depended on the speed and accuracy with which it generated 'daughters'. Subject to a continual risk of disruption by chemical and physical agents in the environment the lifetime of any relatively complex molecule must have been strictly limited. In addition primitive replication was probably highly inaccurate so only a relatively small proportion of copies would have been viable. To avoid extinction a replicator needed to produce at least one viable copy before being disrupted and even so there remained a definite chance of extinction which decreased inversely with the expected number of descendants per generation. It is, therefore, highly unlikely that life began all at once in a particular place at a particular time. Much more plausibly large numbers of colonies of different replicators were appearing and disappearing over the same period of time. Gradually, and inevitably, the colonies became larger as the faster and more accurate replicators outlasted the others. The mean levels of accuracy and speed of replication continued to rise while the freely available raw material for biosynthesis was depleted, leading to intensive interspecies competition and strong selection for catalytic activity and hence, for example, to the evolution of organised cells capable of binary fission.

Coping with errors

At a fundamental level evolutionary survival is the preservation of a dynamic balance between information, or

order, and entropy, or disorder. Indeed, given the capacity for replication, neo-Darwinist natural selection may be seen as the direct result of interaction between the information stored in living organisms and the entropy of their environment. The diversity of present-day species testifies to the wide range of evolutionary strategies that are effective for survival, and it is my intent in this paper to demonstrate why ageing should be one of these.

A vital manifestation of entropy is the error inherent in all processes of macromolecular information transfer within living cells. To sustain the prospect of further evolutionary change and so improve its chance of ultimate survival an organism must make occasional copying errors. Too many mistakes would be a serious handicap so what is needed is a compromise, a tolerable level of error sufficient to allow necessary change. Given enough time natural selection ensures that organisms with optimal error levels will dominate. The presence of any error at all, however, poses interesting problems for our understanding of molecular biology.

Orgel² pointed out that cellular proteins could be divided roughly into two sets, those involved in the transfer of information from DNA, in replication and in protein synthesis, and those involved in other cell functions. The presence of a low level of error in the synthesis of the second set of proteins would have relatively little effect on cell viability. But erroneous synthesis of some of the first set might have serious consequences since abnormal proteins that retained activity but had reduced specificity could produce further defective proteins and so generate a positive feedback of errors leading to a progressive decline in the accuracy of the protein-synthesising machinery and ultimately to a lethal 'error catastrophe'. Orgel subsequently revised³ his original suggestion that the error frequency must increase exponentially and showed that depending on the amount of error feedback allowed, stability of translation may be achieved despite the presence of error. In some circumstances, however, error catastrophe would still occur. Hoffman⁴ has proposed a more detailed model of the feedback of errors in the protein-synthesising machinery and concluded that translation is too stable to allow 'error catastrophe' to play any real part in present-day organisms. He supposes the translation apparatus to comprise a set of polypeptide 'adaptors' responsible for inserting amino-acids into a peptide chain according to the sequence specified by the DNA. The average accuracy of the i th generation of adaptors (assuming these to be in some sense discrete generations)—that is the probability that an adaptor of the i th generation inserts the correct amino-acid—is denoted by q_i . Then what matters is the form of the functional relationship between q_i and q_{i-1} . If $q_i > q_{i-1}$ the error level

*National Institute for Biological Standards and Control, Holly Hill, London NW3 6RB.

Complete accuracy

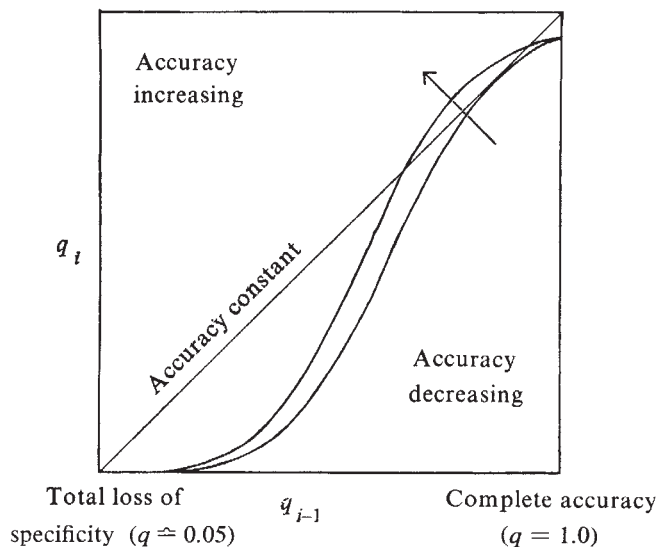


Fig. 1 A simplified model^{4,5} of the feedback of errors in the translation apparatus assumes a set of polypeptide "adaptors" to be responsible for inserting the specified amino acids into peptide chains. The curves illustrate the form of relationship between the average accuracy, q_i , of the generation i adaptors and the average accuracy, q_{i-1} , of the immediately preceding generation. The arrow indicates the direction of displacement of the curve as the average proportion (R) of activity retained by erroneous adaptors decreases.

decreases while if $q_i < q_{i-1}$ it increases. According to Hoffman the dependence of q_i on q_{i-1} is S-shaped with a region where $q_i > q_{i-1}$ for any plausible present-day apparatus. The presence of this stable region forms the basis of Hoffman's conclusion that 'error catastrophe' is not relevant to present-day organisms. Kirkwood and Holliday⁵ point out, however, that Hoffman's model requires that the activity of an erroneous adaptor be reduced to a level so low that feedback of errors is virtually prohibited and argue that a considerable body of biological evidence does not support this assumption. They define a parameter R that represents the proportion of activity retained by erroneous adaptors and show that the value of R determines whether or not stability is attained (see Fig. 1), low values of R resulting in stability (Hoffman's model is equivalent to the case where R is fixed and very small). The introduction of this parameter into the model is of great importance since it allows a degree of freedom on which evolution may work. For example, the development of a 'scavenging' system that would degrade erroneous protein would reduce the value of R .

It should immediately be recognised that this model is a great oversimplification and cannot be viewed as anything more than a broadly descriptive account of the possibilities of error catastrophe or stability. It does, nevertheless, serve to illuminate the general qualitative features that can be expected of the feedback of errors in the protein-synthesising machinery and it forms a useful basis for discussing their relevance to the evolution of present-day organisms.

As primitive cells evolved, each containing its own translation apparatus, some stability of error must have been achieved. Error catastrophes may, of course, have been initiated in some daughter cells but the expected number of 'catastrophe-free' offspring per generation must have been greater than one. Selection would have favoured those cells with the highest proportion of 'catastrophe-free' offspring and so increased the level of stability.

This can be understood quite easily in terms of the Kirkwood/Holliday model. As R (the proportion of activity retained by erroneous adaptors) decreases the S-shaped

curve relating to q_i to q_{i-1} moves steadily upwards and to the left. Thus an R -value can be chosen such that the curve just touches the line $q_i = q_{i-1}$ giving a point of precarious balance. A slight further decrease in R produces a small stable region ($q_i > q_{i-1}$). Random fluctuation in q , however, may allow a jump across the stable region into incipient error catastrophe. Decreasing R decreases the chance of such a jump. Thus we can see that selective pressure continues to reduce R until the chance of jumping across the stable region becomes very small. This would have been particularly true for organisms exposed to environments where the risk of externally introduced errors was high.

On this basis we can expect to find that in present-day organisms the stability of translation will be very high. At this point, however, it is important to draw a distinction between unicellular and multicellular organisms. In discussing ageing our attention is confined to the latter in which the cells may be grouped as somatic cells which do not contribute information to succeeding generations and germ line cells which do. We shall see that there are good reasons why an organism should have evolved a different approach to error regulation in these two categories of cells and that these are sufficient to explain why we age.

The phenomenon of ageing

Since Hayflick and Moorhead⁶ gave the first clear demonstration that human fibroblast cells have a finite lifespan *in vitro* these cells have been studied as a possible model for ageing *in vivo*. Fibroblast lifespan has been shown to be determined largely by cell generations and not by chronological time⁷, and Martin *et al.*⁸ have proved that the finite division potential is related to ageing *in vivo* by demonstrating that culture lifespan declines with donor age. A similar conclusion was indicated by the finding that cells from patients with diseases of premature ageing showed reduced division potential.

Senescent fibroblasts are visibly sick and cannot compete with young cells. It is therefore surprising that cellular selection does not act against the senescent cells leaving the culture permanently viable. The most satisfactory explanation is that cells become "committed" to senescence while still outwardly healthy and that a number of cell generations then elapse before the death of the resultant clone. A model on these lines has been shown to explain the finite lifespan of human fibroblasts in culture⁹. This model is very simple in structure. It is assumed that any cell not yet committed to senescence gives rise to committed daughter cells with some probability P for each daughter and that an incubation period of M cell generations is required before death. If $P \geq 0.5$ the expected number of uncommitted daughters per uncommitted cell is ≤ 1 , so mortality is inevitable. If $P < 0.5$ the uncommitted cells can in principle increase in number. In practice, however, the culture size is strictly limited and surplus cells are discarded with the result that the uncommitted cells are lost 'by dilution' unless the culture size is very large (routine cultures contain 10^6 – 10^7 cells). Experimental results¹⁰ strongly support this model and suggest that $P \approx 0.275$ and $M \approx 55$ – 60 .

The conclusion we might be tempted to draw from the observed mortality of cell cultures, namely that ageing *in vivo* is the result of the body's cells simply running out of division potential, is denied by the data of Martin *et al.*⁸ which show that mean culture lifespan (population doublings) $\approx 42 - 0.2 \times$ donor age so that cells from even 90–100-yr-old individuals are capable on average of 20–25 population doublings. A specific feature of the commitment model is that it predicts a sharp fall in the cell population growth rate as the first committed cells reach the end of their incubation period and die (see Fig. 2) since these constitute approximately a quarter of the population. For normal cell cultures the change of growth rate may be

expected 20–25 population doublings before the end of the culture lifespan. Tissue in old people may therefore be experiencing this fall in growth rate. This conclusion is supported by the estimated value of 55–60 cell generations for the incubation period which allows $\approx 10^{17}$ cells to be produced from a single cell before the change in growth rate would occur, a number reasonably close to the total number of cells likely to be required in a normal human lifetime. Ageing is a complicated process and it may be a mistake to seek too simple an explanation. It is, however, possible that the relatively sudden loss of a substantial proportion of cell replicative potential may account for the physical emaciation and loss of homeostasis associated with senile decay.

Why grow old?

The commitment model closely resembles the model for the feedback of error in the protein synthesising machinery if we equate P with chance of jumping from a stable to an unstable state and M with the metabolic period required for an error catastrophe to result. Fifty-five to sixty cell generations may seem surprisingly long in this context, but q_i initially changes slowly and it can be shown that many generations of adaptors may have to elapse before q_i falls very far.

Experiments show that increasing the error frequency reduces the lifespan of human fibroblasts¹¹ and that old cells contain higher levels of erroneous functional and translational enzymes than young cells^{11,12}. A general error catastrophe thus offers a plausible explanation for the ageing of human fibroblasts. That it offers a likely explanation may be seen from the following argument.

As Hamilton and others have recently emphasised (see ref. 13) natural selection must operate at the level of the replicating unit, the gene. To maximise their chance of survival genes must always replicate between succeeding generations as rapidly and accurately as allowed by their environment. Accuracy can only be achieved through the

expenditure of energy, either by direct proof-reading, kinetic proof-reading as proposed by Hopfield¹⁴, or by destroying erroneous products. The latter may occur either at a molecular level by scavenging for erroneous protein or at a cellular level by means of a 'suicide' protein that kills the cell when the error level begins to rise^{2,15}. Accuracy in the germ line is vital for gene survival but a high level of accuracy in somatic cells may be a luxury our genes do better to forego. Ageing may, therefore, be the result of an energy-saving switching off of the mechanisms responsible for high accuracy in the translation apparatus at or around the time of differentiation of somatic cells from the germ line. This may leave the somatic cells in a relatively unstable state where a small random increase in errors commits them to an eventual error catastrophe.

This strategy is evolutionarily stable in the sense defined by Maynard Smith and Price¹⁶ provided that the risk to the individual of accidental death is not too low since the alternative of maintaining the high accuracy in all cells which might allow immortality would carry the penalty of an increased energy requirement. This penalty would mean that, other things being equal, an individual would grow and reproduce more slowly or would be subject to a greater risk of starvation or accidental death while seeking food.

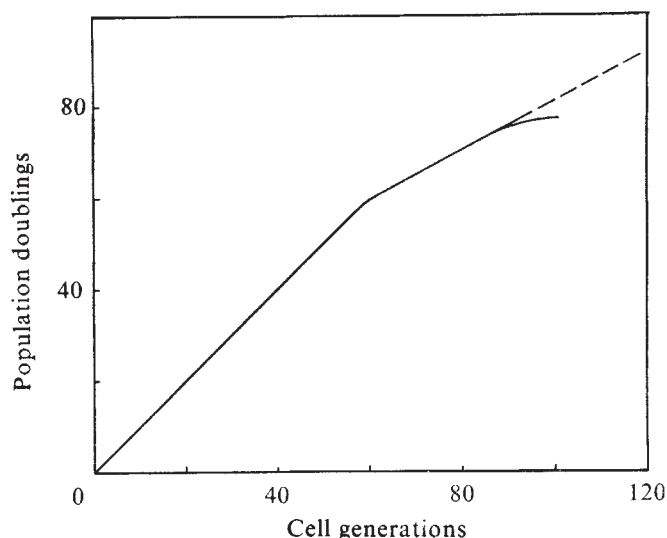
Conclusions and implications

There are three requirements for a satisfactory theory of ageing. It must be theoretically plausible, it must be supported by experimental evidence, and it must make evolutionary sense. The general error theory satisfies all these requirements. The main rival theory suggests that ageing is genetically programmed in much the same way as morphogenesis. Although this is theoretically plausible no specific model has been proposed that is amenable to experimental test and no convincing argument has been put forward to favour the evolution of such a programme. A further theory¹⁷ suggests that ageing is the result of natural selection delaying the time of onset of deleterious genes, so that senescence and death result from the accumulation of late-acting harmful genes. This well-argued theory is, however, either circular or incomplete as it presupposes ageing in a general sense since the time of action of a gene during adulthood is determined not by chronological time but by its biochemical environment. If the time-keeping process is a genetic 'clock' the theory is identical with the programme theory. If it is an increase of errors it may be incorporated as a part of the general error theory. A third possibility, that the genes are triggered by the accumulation (or depletion) of some otherwise innocuous substance, would need to be specifically defined and tested before it could be seriously entertained.

If the argument presented here is correct, we may hope to understand some of the many aspects of ageing by detailed study of the processes of error regulation. We may examine how different species have adapted these processes to meet their individual needs and, in particular, we may hope to gain insight into the abnormalities represented by diseases of premature ageing and by the apparently immortal transformed cells of malignant tumours. It may, for example, be possible that oncogenic viruses interfere with these processes to enhance the accuracy of their own replication in a way that restores some error-committed cells to stability. Such a return to a quasi-germ-line state might explain the reappearance of foetal proteins in malignant tissue. We should remember too that our ageing processes evolved at a time when our average lifespan was considerably shorter, just long enough in fact for us to fulfil our basic evolutionary role of bearing children and nurturing them to reproductive age. In modern technological societies the majority of us reach old age and mortality need not necessarily be evolutionarily stable.

I am very grateful to Dr R. Holliday for continued help

Fig. 2 Predicted fall in cell population growth rate according to a model of the commitment to senescence of diploid human fibroblasts⁹. Cells not yet committed are assumed to give rise to committed daughter cells with probability P (≈ 0.275) for each daughter. M (≈ 55 –60) cell generations after commitment the resultant clone is assumed to die. An initially uncommitted population doubles with each cell generation until, after M generations, the first deaths occur and the population growth rate falls. During routine cell culture, when the population size is strictly limited, uncommitted cells are lost "by dilution" and the population dies out (solid curve). Very large cultures may, however, continue indefinitely at the reduced growth rate (broken line).



and encouragement with this work and to Professor J. Maynard Smith and Dr L. E. Orgel for discussion.

1. Weismann, A. in *Weismann on Heredity* (eds Poulton, E. B., Schönland, S. & Shipley, A. E.) 2nd edn, 23–42 (Oxford University Press, 1891).
2. Orgel, L. E. *Proc. natn. Acad. Sci. U.S.A.* **49**, 517–521 (1963).
3. Orgel, L. E. *Proc. natn. Acad. Sci. U.S.A.* **67**, 1476 (1970).
4. Hoffman, G. W. *J. molec. Biol.* **86**, 349–362 (1974).
5. Kirkwood, T. B. L. & Holliday, R. *J. molec. Biol.* **97**, 257–265 (1975).
6. Hayflick, L. & Moorhead, P. S. *Exp Cell Res.* **25**, 585–621 (1961).

7. Littlefield, J. W. in *Variation, Senescence and Neoplasia in Cultured Somatic Cells*, 62–64 (Harvard University Press, Cambridge, Massachusetts, 1976).
8. Martin, G., Sprague, C. & Epstein, C. *Lab. Invest.* **23**, 86–92 (1970).
9. Kirkwood, T. B. L. & Holliday, R. *J. theor. Biol.* **53**, 481–496 (1975).
10. Holliday, R., Huscchtscha, L. I., Tarrant, G. M. & Kirkwood, T. B. L. *Science* (in the press).
11. Holliday, R. & Tarrant, G. M. *Nature* **238**, 26–30 (1972).
12. Linn, S., Kairis, M. & Holliday, R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2818–2822 (1976).
13. Dawkins, R. in *The Selfish Gene* 13–21 (Oxford University Press, 1976).
14. Hopfield, J. J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4135–4139 (1974).
15. Holliday, R. *Fedn Proc.* **34**, 51–55 (1975).
16. Maynard Smith, J. & Price, G. R. *Nature* **246**, 15–18 (1973).
17. Medawar, P. B. *An Unsolved Problem in Biology* (H. K. Lewis, London, 1952).

articles

Drift of the major continental blocks since the Devonian

E. Irving

Centre Géologique et Géophysique, Université des Sciences et Techniques du Languedoc, 34060 Montpellier, France, & Earth Physics Branch, Department of Energy, Mines and Resources, Ottawa, Canada

Palaeomagnetic evidence indicates that the continents have been in a more-or-less continuous relative motion. At the end of the Palaeozoic there was a redistribution of the major continental blocks that occurred without the formation of new ocean between them. Wegener's Pangaea seems only to have lasted a few tens of millions of years.

MAPS summarising the drift of the major continental blocks since the middle Devonian (–375 Myr ago) are presented here. They are based mainly on a global synthesis of the palaeomagnetic data¹, including the most recent results from the USSR². Reasonably reliable apparent polar wandering (APW) paths for the major continental blocks can now be drawn for the Middle Carboniferous (325 Myr ago) onwards, and are given in Figs 1 to 3. They have been obtained by first referring all the palaeomagnetic poles to a scale of Myr³. The oldest poles from each continental block were then grouped into an interval of 40 Myr duration, and an average calculated to obtain the oldest point. The limits of the interval were then moved forward by 10 Myr and a second average taken, and so on. Only every fourth value is independent. It would be desirable to use averages over a shorter interval, but as yet the data are neither numerous enough, nor sufficiently well spread in time. Calculations using 30 and 40 Myr intervals yielded almost identical mean values, but at 20 Myr more scatter was evident. The mean 95% errors are 6° (North America), 7° (Eurasia) and 7° (Gondwana).

The calculations were made assuming that the geomagnetic field averaged to an axial geocentric dipole (AGD). There is evidence that even when averaged over millions of years, the field may have had long-term non-dipole components⁴ sufficient to cause polar errors of about 5°, and there are physical reasons why this might be so⁵. Analysis⁶ of the geomagnetic field over the past 25 Myr yields no evidence for departures greater than about 5°. Indeed, it has recently been shown that during the Phanerozoic there is no significant departure of the field from that of an AGD within the errors in the palaeomagnetic determinations⁷. There is good agreement between the palaeomagnetically determined latitude and the palaeoclimatic evidence, and this supports the AGD assumption in a general

way⁸; for example, glaciations occur in Australia from the Middle-Carboniferous to the end of the Permian⁹, and their disappearance coincides with the motion of the pole away from Australia (Fig. 3). Finally, there is excellent internal agreement among Gondwana poles from localities spaced over 9,000 km, among poles from localities spread over 5,000 km in North America, and among poles spread over 6,000 km from Eurasia; note, for example, the high precision ($k = 103$ to 406) of Gondwana poles in the interval 270 to 290 Myr ago from Argentina, Morocco, Madagascar, Tanzania and Australia (Table 1). It is unlikely that such internal agreement would have been observed within blocks that at the time were adjacent to one another, if the field really had had very large non-dipole components. Therefore there are grounds for supposing that the poles of Table 1 provide estimates of the geographical pole with an average accuracy of 6° or 7° ($P = 0.05$).

Since the advent of plate tectonics, maps showing displacements of the continents have been made using evidence mainly from the oceans, information that extends back only to the early Jurassic. Such maps indicate that at about 200 Myr ago the continental crust was grouped into a single supercontinent¹⁰, referred to here as Pangaea A. Pangaea A is essentially Wegener's reconstruction¹¹. But for how long, before 200 Myr ago, did Pangaea A exist? This may be studied by comparing the latitudes for reference localities that were once adjacent but are now far apart using the paths of Figs 1 to 3. If Pangaea A existed the latitudes should be the same. For the comparison between North America and Africa (Fig. 4) the reference locality used is midway between Cape Blanc (locality L 21°N, 19°W in the African frame) and Cape Hatteras (locality M 37°N, 74°W in the North American frame) using the most recent reconstruction¹² of Pangaea A; $\Delta\lambda = \lambda_L - \lambda_M$, and λ has been calculated from Table 1 using equation (9.1) of ref. 8. For the Lower Jurassic (190–160 Myr) $\Delta\lambda$ is effectively zero, so that the palaeomagnetically determined latitudes agree with Pangaea A. Before that $\Delta\lambda$ is positive and significant, suggesting that during the Permian and much of the Triassic, Africa had a more northerly position relative to North America than that indicated by Pangaea A. This result is independent of the choice of reference localities, and is essentially the same for other reconstructions^{13,14}. This discrepancy, although not previously so well documented, has been known for many years, and

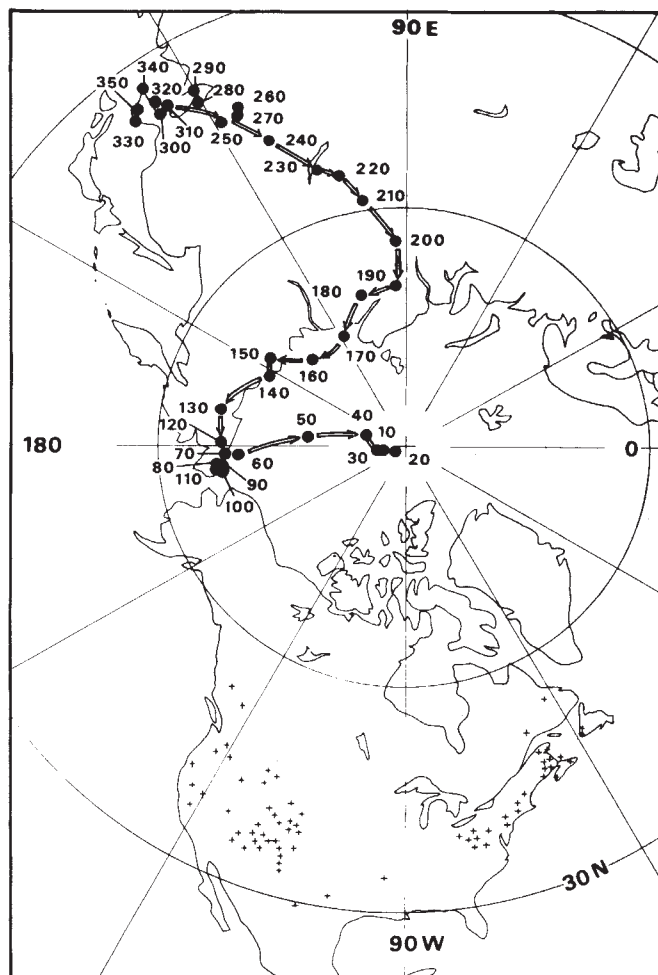


Fig. 1 Apparent polar wander (APW) relative to North America since the Devonian. Sampling localities are marked by crosses. Errors listed in Table 1.

several workers¹⁵⁻¹⁸ have proposed that there was a large post-Triassic shear along Tethys (the Tethys twist¹⁷), that was supposed to account for the Alpine deformation^{16,17}. This idea is no longer acceptable because the early Jurassic palaeomagnetic results are now known to be in excellent agreement with Pangaea A (Fig. 4); the motions needed to account for the discrepancy must have occurred before the Alpine Orogeny, and before the creation of the western part of Tethys in the late Triassic. For the comparison between North America and Europe the corresponding $\Delta\lambda$ values using Orphan Knoll as the reference locality (P 50°N, 47°W in the North American frame, and Q 50°N, 17°W in the European frame $\Delta\lambda = \lambda_P - \lambda_Q$) do not differ from zero in the Permian and lower Jurassic, but there is a significant departure of about 10° in the Triassic (Fig. 4). It has been proposed¹⁹ that this is best explained by assuming that Eurasia moved southward with respect to North America during the lower and middle Triassic. Although this is probably the correct interpretation of the discrepancy itself, it alone does not explain the remarkable fact that later, in the lower Jurassic, the results are again in excellent agreement with Pangaea A (Fig. 4). Tectonic explanations for the discrepancies of Fig. 4 that are globally compatible with all the available palaeomagnetic data are now presented.

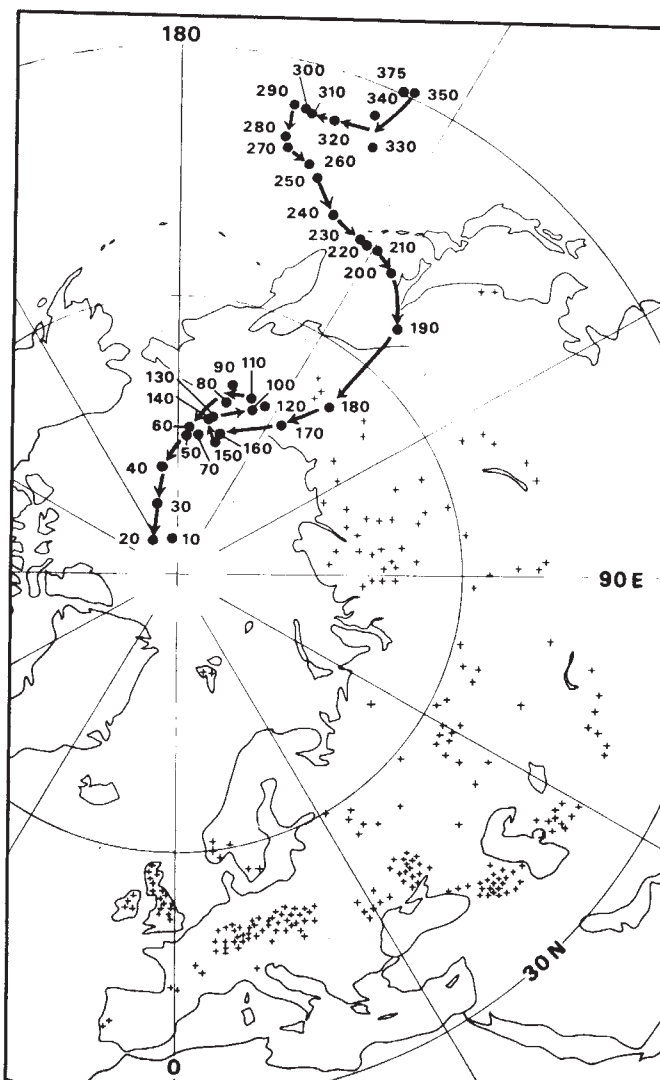
APW paths define the latitudes and azimuths but not the longitudes of continents. In order to obtain maps for post-Pangaea A time (200 Myr ago) the latitudes of the continents have been fixed palaeomagnetically, and their relative longitudes have, for the most part, been taken from plate tectonic reconstructions^{10,20-22}. Maps for times before 200 Myr ago have

been made by working backwards, interval by interval, from Pangaea A, maintaining consistency with the palaeomagnetically determined latitudes and azimuths, but restricting relative motions to a minimum and disallowing major overlap between continents.

Description at maps

The suggested reassembly for the mid-late Devonian (375 Myr ago) is shown in Fig. 5a. In the north there is a more-or-less coherent supercontinent, Laurasia, consisting of the Laurentia, Baltic-Russian and Siberian cratons and their bordering Early Palaeozoic foldbelts. Several authors^{27,28} have argued that in the Devonian there was a wide ocean on the site of the Ural foldbelt. But, the palaeomagnetic poles from the Devonian of the Russian and Siberian platforms, are in reasonable agreement, and therefore it is assumed that there was no more than a narrow trough between them. The mean Devonian pole for the Russian platform is 32°N, 161°E, $\alpha_{95} = 5^\circ$ (5.120), for the Dneister region it is 39°N, 161°E, $\alpha_{95} = 7^\circ$ (5.130), for the Ural region it is 38°N, 165°E, $\alpha_{95} = 9^\circ$ (5.177) and for the Siberian platform 24°N, 149°E, $\alpha_{95} = 8^\circ$ (5.178). (The numbers in brackets refer to entries in ref. 1 where the bibliography and analytical details may be found.) Because of the longitude indeterminacy this solution is not a rigorous one, but only the simplest, which devolves from the constraint of minimum

Fig. 2 APW relative to northern Eurasia since the Devonian. See legend Fig. 1.



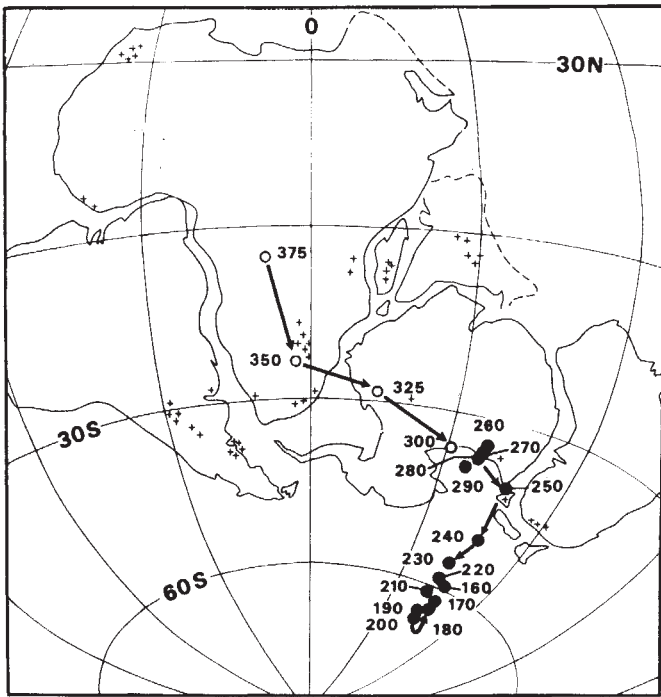


Fig. 3 APW relative to Gondwana for the Late Palaeozoic and Early Mesozoic. ●, Are the poles of Table 1 and have been obtained by statistical analysis; ○, for the interval 375 to 325 Myr ago, are approximate poles, obtained by visual inspection of sparse data from Australia and South America only. This interpretation follows that given in refs 43, 45 and 46, and is speculative. Gondwana began to split in the Jurassic, and an average path cannot be obtained for later times. See legend Fig. 1.

motion. If the Urals are the site of a collision between the two platforms it seems more likely that the major movement occurred in the early rather than the late Palaeozoic. This problem is not of direct interest in the present context and will be considered again later (Morel and Irving, in preparation). In the late Devonian assembly of Fig. 5a Spitzbergen adjoins Ellesmere Island and Greenland²³, and northern Alaska is rotated with respect to North America^{24,25}. There is at present no palaeomagnetic information from the northern two-thirds of Alaska that would shed light on this speculation, but there is indirect evidence. In the Mid-Cretaceous (100 Myr ago) the pole positions relative to Eurasia and North America make it impossible to place them together without substantial overlap of their continental shelves (Figs 1 and 2). Alaska, or NE Siberia, or both have moved. Chukotsk²⁶, with or without the Kolyma block (Irving and Geur, in preparation), may have been continuous with northern Alaska. It is possible that all three elements (northern Alaska, Chukotsk and Kolyma) moved together, their tectonic juncture with Asia being the huge fault systems of the Khrebet Cherskogo. Alternatively, Kolyma may have been welded on to Asia separately³⁴, and this is accepted here. The possible position of NE Siberia relative to northern Alaska is indicated by *H?* in Fig. 5. A further feature of Fig. 5a to i is an indentation in the northern border of Laurasia. This is the Sinus Borealis, or proto-Arctic, which closed in the early Cretaceous (see below). To the south of Laurasia there may have been an ocean about 2,000–3,000 km wide (see legend to Fig. 5). The southern continents are assembled into a supercontinent, Gondwana²⁹.

In the zone of interaction between Laurasia and Gondwana, the positions of many crustal elements are uncertain. For example, Spain and central Europe may have been situated in position *B* or *C* (see ref. 30) of Fig. 5, and the USA south and east of the Appalachians could have occupied either positions

A or *B*. The best recent data³¹ favour position *B* for southern France. The data from Spain are indecisive. Nor is very much known about the positions of Iran, China³² and southern Indo-China. This paper is concerned with the major continental blocks and no detailed consideration of these questions is given.

At the Devonian–Carboniferous boundary, Laurasia and Gondwana are converging and they meet in the Mid-Carboniferous. The relative longitudes of Fig. 5c were chosen so as to minimise their subsequent relative motion. Africa is opposite

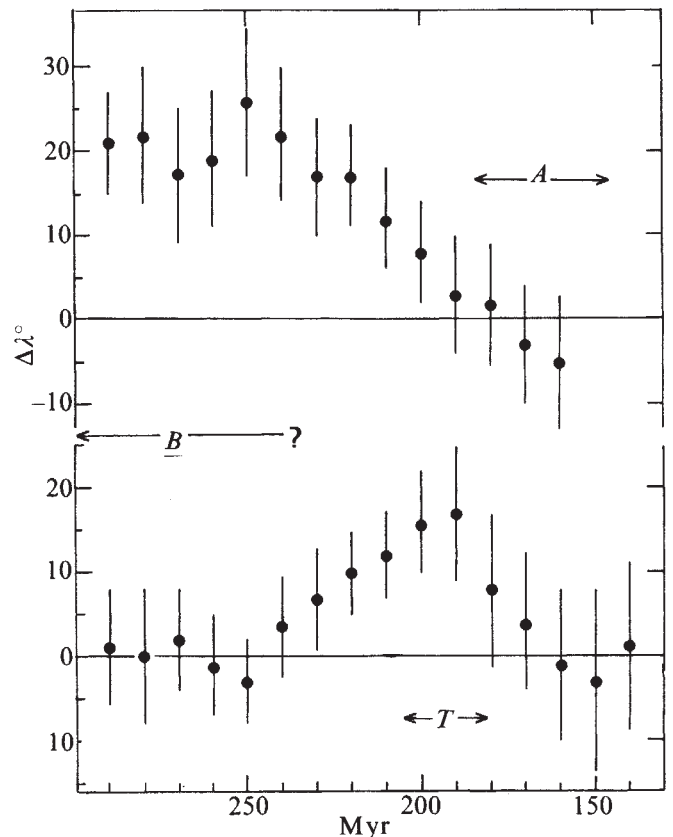


Fig. 4 Latitude differences ($\Delta\lambda$) for reference localities from North America and Gondwana (above), and North America and Europe (below), for the late Palaeozoic and Early Mesozoic using the reconstruction of Pangaea B¹². See text for method of calculation. The errors at ($P = 0.05$) are given. The approximate durations of Pangaea A and B are marked. *T* is the suggested time of opening of the western Tethys.

Europe, and South America opposite North America. This is referred to as Pangaea B. A similar reconstruction for the Permian has recently been given independently by Westphal³⁴, and the earliest record of it is in ref. 9 (Fig. 10, 17). There are, however, no palaeomagnetic constraints on longitude, so Gondwana could have been further east, or it could have been situated west of North America. These optional configurations^{17,18} are considered less likely because the subsequent motions needed to achieve the known late Jurassic configuration (Fig. 5i) would have been greater—a factor of two greater in the latter instance. In Gondwana the Samfrau²⁹ superfoldbelt develops. It extends for 6,000 km from the Tasman foldbelt, through the Ellsworth and Cape foldbelts to the central Andes. By the end of the Carboniferous the fit between Laurasia and Gondwana tightens, and apparently remains unchanged in the early Permian, although small adjustments, undetectable palaeomagnetically, could be occurring (Fig. 5d to f).

At the Permian-Triassic boundary the palaeomagnetically determined latitudes now allow Gondwana to rotate anticlockwise, and the transformation from Pangaea *B* to *A* to begin. This is made possible by a northward movement of North America¹⁹ along a zone (*XM*), located, it is suggested, between Ireland and Newfoundland, and passing north to the Sinus Borealis. By the late Triassic the rotation of Gondwana

Table 1 Average poles for the major continental blocks

<i>T</i>	<i>W</i>	<i>N(R)k</i>	λ, ϕ	α_{95}
North American poles (°N, °E)				
20	20	10(09.95)175	88,171	4
30	20	10(09.89)82	86,152	5
40	20	07(06.91)64	85,167	8
50	40	12(11.76)45	78,173	6
60	40	11(10.81)51	73,186	6
70	40	11(10.87)79	70,187	5
80	40	09(08.91)85	69,192	6
90	40	08(07.93)95	69,194	6
100	40	05(04.99)458	69,190	4
110	40	04(03.99)503	67,186	4
120	40	05(04.92)52	68,179	10
130	40	06(05.89)44	67,169	10
140	40	07(06.75)24	72,152	13
150	40	08(07.77)31	72,147	10
160	40	08(07.83)40	76,131	9
170	40	07(06.87)45	76,114	9
180	40	09(08.80)41	72,106	13
190	40	07(06.91)64	70,094	8
200	40	07(06.96)134	64,093	5
210	40	11(10.84)61	59,100	6
220	40	07(06.92)79	55,103	7
230	40	08(07.81)37	53,108	9
240	40	09(08.84)50	49,114	7
250	40	06(05.94)82	43,120	7
260	40	09(08.79)38	42,115	9
270	40	10(09.78)40	43,117	8
280	40	08(07.77)31	39,121	10
290	40	11(10.72)36	38,121	8
300	40	10(09.90)90	38,126	5
310	40	10(09.91)95	37,124	5
320	40	11(10.86)72	36,126	5
330	40	07(06.93)85	36,129	7
340	40	06(05.93)73	33,126	8
350	40	05(04.92)50	35,128	10
Poles for northern Eurasia (°N, °E)				
20	30	51(49.90)31	85,198	4
30	30	30(29.87)26	81,190	5
40	40	31(30.08)33	78,180	5
50	40	29(28.25)37	76,174	4
60	40	28(27.28)37	75,175	5
70	40	24(23.37)40	75,170	5
80	40	16(15.58)27	71,166	6
90	40	16(15.57)30	69,164	6
100	40	17(16.59)37	71,157	6
110	40	15(14.75)57	70,159	5
120	40	15(14.70)46	70,153	6
130	40	15(14.48)27	74,166	7
140	40	12(11.59)27	73,170	8
150	40	13(12.49)24	76,162	9
160	40	13(12.48)23	76,163	9
170	40	12(11.65)31	71,146	8
180	40	12(11.55)25	65,139	9
190	40	10(09.72)33	54,140	9
200	40	25(24.14)28	49,147	6
210	40	41(39.91)37	49,149	4
220	40	51(49.56)35	48,153	3
230	40	50(48.60)35	48,153	3
240	40	43(42.11)47	47,157	3
250	40	27(27.46)49	45,163	4
260	40	17(16.79)77	42,165	4
270	40	21(20.69)65	42,168	4
280	40	16(15.65)43	40,166	6
290	40	20(19.48)36	37,167	5
300	40	20(19.33)29	36,166	6
310	40	17(16.34)24	36,165	7
320	40	20(18.99)19	36,162	8
330	40	14(13.29)18	38,157	10
340	40	14(13.20)16	35,158	10
350	40	12(11.31)16	29,156	11
375	40	08(07.93)95	30,155	6
Poles from Gondwana (°S, °E)				
160	30	08(07.89)63	60,076	7
170	30	12(11.80)53	63,076	6
180	30	12(11.84)70	64,077	5
190	30	09(08.86)57	67,074	7
200	30	07(06.92)80	67,076	7
210	30	10(09.84)56	68,072	6
220	30	11(10.85)69	58,072	6
230	40	12(11.56)25	56,072	6
240	40	08(07.86)51	51,077	8
250	40	05(04.90)40	41,076	12
260	40	05(04.96)98	36,068	8
270	40	05(04.96)103	37,067	8
280	40	06(05.96)118	38,068	6
290	40	05(04.99)406	40,065	4

T is geological time in Myr. *W* is the width of the averaging interval in Myr. *N* is the number of values used, *R* the resultant and *k* the precision⁴⁸. λ, ϕ are the latitude and longitude of the mean pole. α_{95} is the 95% error. The data used are the starred entries in the Ottawa catalogues, and omitting results from deformed beds in foldbelts; for example, the selection for North America is essentially the same as that of ref. 50. Results from Speckled Sandstone of India⁵¹ are also omitted. A list of the data used may be obtained from the author. The values for Gondwana have been rotated relative to Africa using the reconstruction of ref. 52.

relative to Laurasia stops, and the motion along *XM* reverses (Fig. 5*h*). North America is now locked to Gondwana, and both rotated anticlockwise relative to Europe, opening the western Tethys and forming Pangaea *A*. Mexico is assumed to move NW relative to North America³³. During the transformation from Pangaea *B* to *A* it is as if Gondwana forces North America northwards out of its path, and then, carrying North America along with it, opens Tethys. The total relative motion along the Hercynian super-foldbelt is about 3,500 km. Movement is completed by the late Triassic, but its start is poorly determined. Some shear may have begun in the early Permian, and the time required may have been as long as 75 Myr (average relative velocity 5 cm yr⁻¹). The northward movement of North America relative to Europe is at least 1,000 km, although the errors are such that the movement may be as much as 2,000 km (ref. 19). The succeeding southward movement of North America relative to Europe is of the same magnitude. The motion along *XM* is unlikely to be pure shear but also partly extensional, and one of the consequences might be the development of grabens, which, in the North Sea, contain oil. The formation of the early Triassic Siberian Traps, the dextral rotation of Novaya Zemlya, and the opening of the Kara Sea²⁵, may also be caused by it. During the late Triassic, the Cordilleran super-foldbelt apparently develops as a continent-ocean interaction, and extends over 10,000 km from New Zealand to Alaska.

By the late Triassic or early Jurassic Wegener's Pangaea (Pangaea *A*) is achieved (Fig. 5*i*). By the mid-late Jurassic the North Atlantic begins to open, Mexico is able to achieve its present position, and a rift appears between eastern and western Gondwana. In the early Cretaceous, Sinus Borealis closes (Fig. 5*k*), and Spain begins to rotate anticlockwise. Rifting is initiated in the South Atlantic. In the Middle Cretaceous, Tethys is open to the Atlantic (Fig. 5*l*). The Kolyma Block, and surrounding Mesozoic foldbelts, could have been welded onto Eurasia^{27,34} at this time. The region around the north pole is entirely covered with continental crust. During the late Cretaceous the Labrador Sea (*LS*) opens (Fig. 5*m*), Africa moves eastward relative to Europe, and Tethys closes to the west.

At the beginning of the Tertiary, NE Eurasia moves away from North America, allowing northern Alaska and parts of NE Siberia to rotate anticlockwise and open the Canada Basin (Fig. 5*n*); the Canada Basin could therefore be a back-arc basin formed by the migration of a thin sliver of continental crust over the NW Pacific subduction zone. India moves northward, and the Alpine-Himalayan super-foldbelt forms. By the mid-Tertiary, Alaska completes its rotation (Fig. 5*o*), the Atlantic rift system is fully developed, and the Alpine-Himalayan super-foldbelt is established in its present form.

Discussion

The maps are presented as an explanation of the palaeomagnetic record. They have been constructed by adhering to the AGD assumption, within the statistical errors, that are within the limits for non-dipolar fields as they are presently known⁶. The maps indicate that Wegener's Pangaea lasted only a few tens of Myr, and that in the Permian and Mesozoic major continental blocks of North America, Eurasia and Gondwana were in more-or-less continuous motion relative to one another. Other workers, however, have taken a less mobilistic view, and their maps differ in two important respects. Firstly their maps⁴⁹ show the Arctic to be a static region throughout the Mesozoic, the bordering continents of North America and Eurasia being fixed relative to one another. But geologically the Arctic is far from static. There are foldbelts, rifts, subsiding basins and thick igneous sequences, that in other contexts would be taken as evidence of plate motions. Secondly, in most contemporary maps Wegener's Pangaea is regarded as inviolate in the Permian and Triassic, and the palaeomagnetic discrepancies of Fig. 4 are attributed to the presence of large non-dipole components in the geomagnetic field^{28,35,36}, so that the palaeomagnetic pole

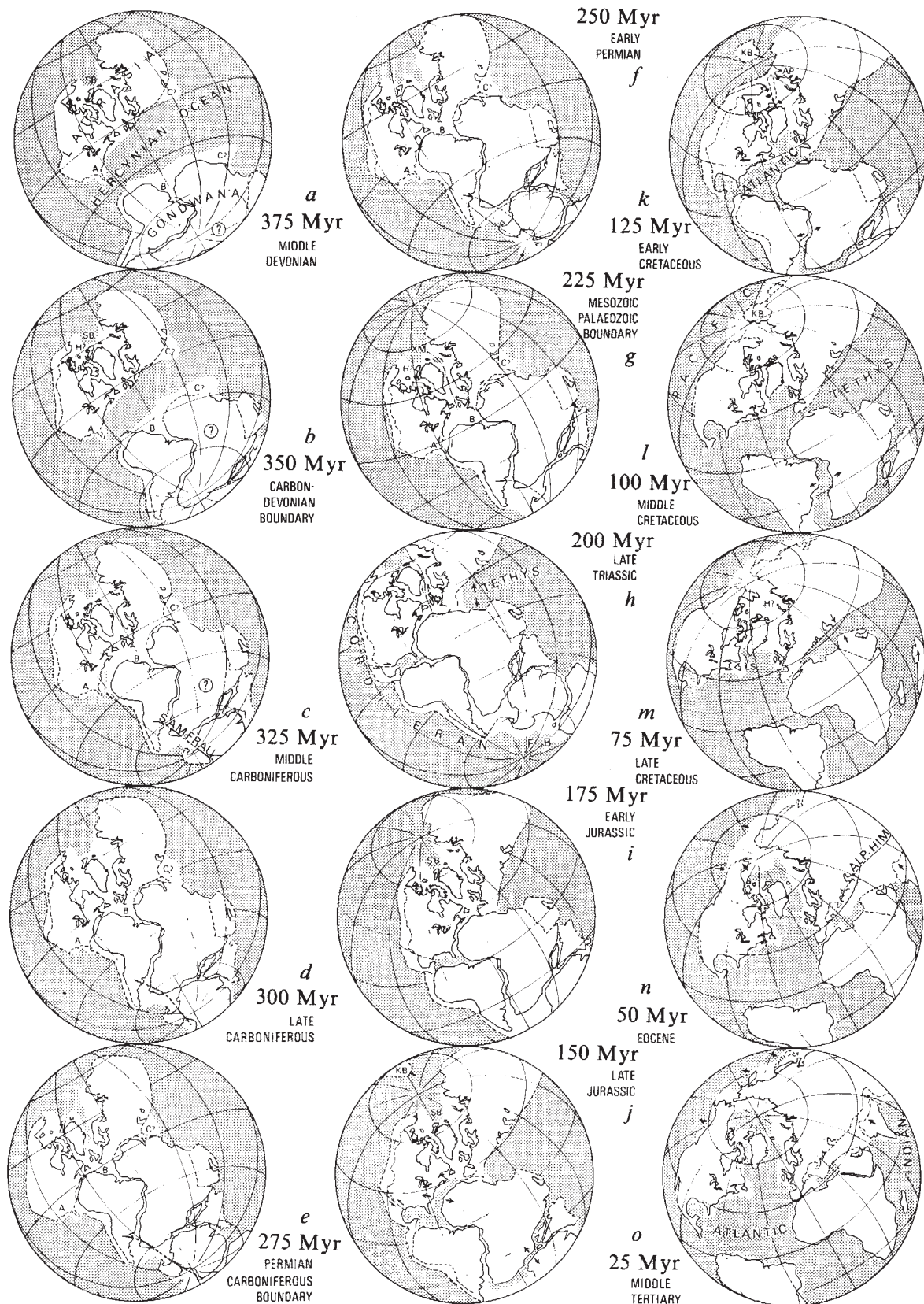


Fig. 5 Continental drift since the Devonian. The latitudes of the major continents have been fixed by placing the palaeomagnetic poles (within their 95% error circles) at the poles of the projection using Geuer's Terrascope¹⁷. This rule is relaxed for *g* and *h* because it is not yet possible to provide satisfactory maps for the *B* to *A* transformation; the necessary details are unlikely to be revealed until a shorter averaging interval can be used, perhaps 20 Myr. *A*, *B* and *C* signify possible positions for micro-continents between Laurasia and Gondwana (see text); *AP* possible zone of subduction during closure of Sinus Borealis *SB*; *H*? possible position of the Chukotsk peninsula relative to Alaska; *KB* Kolyma Block; *LS* Labrador Sea; *XM* suggested location of Triassic shear zone between Europe and North America. The question-marks in Africa in *a*, *b* and *c* draw attention to the possible error in latitude determinations for Gondwana for the Devonian and Early Carboniferous, that arises from the uncertainty in the Devonian and Early Carboniferous pole path for Gondwana (Fig. 3). The width, indeed the existence of the 'Hercynian Ocean', is uncertain. This figure is an attempt to map the movements of the major continent blocks relative to one another and to the pole. No attempt is made to map the positions of minor blocks, and the boundaries of some of the major blocks (for example, North-East Gondwana, and South-East Asia) are highly speculative.

determinations, even when derived by averaging results from many rock units distributed over a large continental block, would be in error by as much as 20°. There is, however, excellent agreement between the palaeomagnetically determined latitudes and plate tectonic reconstructions for the Lower Jurassic onwards (Fig. 5i to o), so that if such large non-dipole components were present, they seem to have disappeared at the beginning of the Jurassic. There is no other evidence for such a change. Although the hypotheses of an inviolate Pangaea and large non-dipole fields in the late Palaeozoic and Mesozoic are still tenable, it is argued here that the alternative tectonic explanations of Fig. 5 are more reasonable; it would be a coincidence if the non-dipole fields were of exactly the form to produce the apparent latitude differences between North America and Eurasia at just the right time to allow the transformation from Pangaea *B* to *A* to occur, followed by the creation of Tethys when required by the geological evidence.

The concept of a mobile Pangaea provides a framework within which to discuss geological events. The Hercynian–Appalachian super-foldbelt could have developed as a result of the collision of Laurasia and Gondwana³⁰. It is cut by large right-lateral wrench faults^{37,38} and it has been suggested³⁷ that they were caused by the relative right-lateral displacement of Laurasia with respect to Gondwana; the displacement is estimated to be 600–1,000 km, and to have occurred between 290 and 250 Myr ago. It post-dates the main Hercynian Orogeny. The sense of displacement is the same, but it is three times less, and apparently began about 40 Myr earlier, than that of Fig. 5. Therefore there could be two diastrophic phases, the smaller earlier displacement of ref. 37 being followed by the larger displacement described here. The two phases could have overlapped in time. The earlier displacements at first met much resistance, and hence were slow and small, and were responsible for most of the observable geological strains. Once a zone of weakness was established, rapid displacements occurred, accompanied by less stress, and fewer observable geological effects. The location of the zone of weakness is not known, but it may have been between Spain and Morocco, and the Brevard Zone of the southern Appalachians is also a possible site. Alternatively, the dating of either the palaeomagnetic results, or the movements along the wrench faults, may be in error. Such large systematic errors are possible, but unlikely. Although the maps of Fig. 5 provide the framework for the discussion of the western part of the Hercynian super-foldbelt between Laurasia and Gondwana, they do not do so for the eastern part situated in central and eastern Asia. It is perhaps worthy of note that palaeomagnetic results from Malaya³² and China¹ indicate that these blocks were in latitudes of 10 to 20°N during the Permian, and therefore they could have been situated in the eastern embayment between Africa and Asia (Fig. 5e). They could have formed part of one or more blocks that interacted with the southern edge of the Siberian craton to form the eastern part of the Hercynian super-foldbelt.

Pangaea *B* existed from the middle of the Carboniferous to the end of the Palaeozoic, about 80 or 90 Myr. It presumably indicates a special type of mantle convection at that time. It corresponds in time to the late Palaeozoic reversed interval (the Kiaman of ref. 39) when changes in the polarity of the geomagnetic field were very infrequent. The significance of this correlation can only be guessed. The onset of the *B* to *A* transformation coincides, within the accuracy of the age estimates, with the onset of frequent reversals at the end of the Permian. This first-order change in the geomagnetic field was probably a response to topographic or temperature changes at the core–mantle interface, and hence to changes in the lower mantle^{40,53}. The *B* to *A* transformation differs from other known Phanerozoic drift episodes in that it does not seem to involve the creation of spreading ridges and new oceans between the major continental blocks. The absence of known spreading ridges is a possible cause of the worldwide regression at the end of the Permian. Exposure of the continental shelves, which house a high proportion of biological activity, could have

been the major environmental change responsible for the widespread extinctions at the end of the Palaeozoic⁴¹. The changes that occurred in mantle convection at the end of the Palaeozoic could have been responsible for both the sudden increase in reversal frequency, and the *B* to *A* transformation, and the latter, in its turn, caused the late Permian extinctions. The identification, made here, of a first-order episode of continental drift, that commenced at the very end of the Palaeozoic or earliest Triassic, may therefore be the major cause of the geological break between the Palaeozoic and Mesozoic. This drift episode, whose main feature was a dextral shear between the northern and southern continents about the palaeoequator, may also reflect a major change in mantle convection which produced what is perhaps the most important change in the geomagnetic field yet recorded.

Finally, the geological evidence indicates that during the late Palaeozoic there was apparently no ocean between Europe and Gondwana⁴², and Tethys did not open until the early Mesozoic (Fig. 5g and h agree well with this evidence).

I thank F. Arthaud, R. Black, F. Brun, J. Geuer, J. Hastie, P. Louis, M. Mattauer, P. Matte, D. F. Strong, J. Sweeney, P. Tapponier and K. Whitham, for their help.

Note added in proof: Recent palaeomagnetic studies in the Brooks Range have shown the anticlockwise rotation of northern Alaska⁵⁵, thus confirming the indirect arguments given above.

Received 10 May; accepted 15 August 1977.

- Irving, E., Tanczyk, E. & Hastie, J. *Geomag. Scr. Earth Phys.* nos. 5, 6 & 10 (Branch, E.M.R. Ottawa, 1976).
- Khranov, A. N. *Acad. Sci. USSR Sov. Geophys. Commun. WDC* 3, 1–43 (1975).
- Harland, W. B., Smith, A. G. & Wilcock, B. J. *geol. Soc. Lond.* 120, 250 (1964).
- Wilson, R. L. *Geophys. J. R. Astr. Soc.* 19, 417–437 (1970).
- Hide, R. *Science* 157, 55–56 (1967).
- Wilson, R. L. & McElhinny, M. W. *Geophys. J. R. Astr. Soc.* 39, 570–586 (1974).
- Evans, M. E. *Nature* 262, 676–677 (1976).
- Irving, E. *Paleomagnetism* 1–399 (Wiley, New York, 1964).
- Crowell, J. C. & Frakes, L. A. *J. geol. Soc. Aust.* 17, 115–155 (1971).
- Phillips, J. D. & Forsyth, D. *Geol. Soc. Am. Bull.* 83, 1579–1600 (1972).
- Wegener, A. *Origin of Continents and Oceans* (Methuen, London, 1924).
- Le Pichon, X., Sibuet, J. C. & Francheteau, J. *Tectonophysics* (in the press).
- Bullard, E. C., Everett, J. E. & Smith, A. G. *Phil. Trans. R. Soc. A258*, 41–51 (1965).
- Van der Voo, R. & French, R. B. *Earth Sci. Rev.* 10, 99–119 (1974).
- Nairn, A. E. M. *Overseas Geol. Min. Res.* 9, 302–320 (1964).
- De Boer, J. J. *geophys. Res.* 70, 931–944 (1965).
- Van Hiltten, D. *Tectonophysics* 1, 3–71 (1964).
- Irving, E. *Systematics Ass. Spec. Publ.* no. 7, 59–76 (1967).
- Roy, J. L. *Earth planet. Sci. Lett.* 14, 103–114 (1972).
- Pitman, W. C. & Talwani, M. *Geol. Soc. Am. Bull.* 83, 619–640 (1972).
- McKenzie, D. & Sclater, J. G. *Geophys. J. R. Astr. Soc.* 25, 437–528 (1971).
- Francheteau, J. *Implications of Continental Drift to the Earth Sciences*, (eds Tarling, D. H. & Runcorn, S. K.) 197–202 (Academic, New York 1973).
- Harland, W. B. *Arctic Geology* (ed. Pitcher, M. G.) *Am. Ass. Petrol. Geol.* 599–608 (1973).
- Tailleur, I. L. *Arctic Geology* (ed. Pitcher, M. G.) *Am. Ass. Petrol. Geol.* 526–535 (1973).
- Carey, S. W. *Continental Drift Symp.* 177–358 (University Tasmania, Hobart, 1958).
- Churkin, M. *Arctic Geology* (ed. Pitcher, M. G.) *Am. Ass. Petrol. Geol.* 485–499 (1973).
- Hamilton, W. *Geol. Soc. Am. Bull.* 81, 2553–2576 (1970).
- Smith, A. G., Briden, J. C. & Drewry, G. E. *Palaeont. Ass. Spec. Paper* 12, 1–42 (1973).
- Du Toit, A. L. *Our Wandering Continents* 1–366 (Oliver and Boyd, Edinburgh, 1937).
- McKerrow, W. S. & Zeigler, A. M. *Nature phys. Sci.* 240, 92–93 (1972).
- Zijderveld, J. D. A. thesis, Univ. Utrecht (1975).
- McElhinny, M. W., Haile, N. S. & Crawford, A. R. *Nature* 252, 641–645 (1974).
- Van der Voo, F., Mauk, F. J. & French, R. B. *Geology* 2, 177–178 (1976).
- McElhinny, M. W. *Implications of Continental Drift to the Earth Sciences* (eds Tarling, D. W. & Runcorn, S. K.) 77–85 (Academic, New York, 1973).
- Creer, K. M. *Applications of Continental Drift to the Earth Sciences* (eds Tarling, D. H. & Runcorn, S. K.) 47–76 (Academic, New York, 1973).
- Briden, J. C., Smith, A. C. & Sallomy, J. T. *Geophys. J. R. Astr. Soc.* 23, 101–117 (1971).
- Arthaud, F. & Matte, P. *Geol. Soc. Am. Bull.* (in the press); *Tectonophysics* 25, 139–171 (1975).
- Mattauer, M., Proust, F. & Tapponier, P. *Nature* 237, 160–162 (1972).
- Irving, E. & Parry, L. G. *Geophys. J. R. Astr. Soc.* 7, 395–411 (1963).
- Irving, E. & Pullaiah, G. *Earth Sci. Rev.* 12, 35–64 (1976).
- Schopf, T. J. M. *J. Geol.* 82, 129–143 (1974).
- Argyriadis, I. *Bull. Soc. Geol. France* 17, 56–62 (1975).
- Wensink, H. *Progress in Geodynamics* 190–208 (R. Neth. Acad. Arts. Sci., 1975).
- Creer, K. M. *Earth Sci. Rev.* 6, 369–466 (1970).
- McElhinny, M. W. & Embleton, B. J. *J. Tectonophysics* 22, 1–29 (1974).
- Hailwood, E. A. *Earth planet. Sci. Lett.* 23, 376–386 (1974).
- Geuer, J. W. *Can. J. Earth Sci.* 10, 1164–1175 (1973).
- Fisher, R. A. *Proc. R. Soc. A217*, 295–305 (1953).
- Smith, A. G. & Briden, J. C. *Palaeocontinental Maps for the Mesozoic and Cenozoic* 63 (Cambridge University Press, Cambridge, 1977).
- Beck, M. E. *Am. J. Sci.* 276, 694–712 (1976).
- Wensink, H. *Tectonophysics* 26, 281–292 (1975).
- Smith, A. G. & Hallam, A. *Nature* 225, 139–144 (1970).
- Cox, A. *Rev. Geophys. Space Phys.* 13, 35–51 (1975).
- Westphal, M. *Nature* 267, 136–137 (1977).
- Newman, G. W., Mull, C. G. & Watkins, N. D. *Geol. Soc. Alaska, Abstracts, Anchorage meeting*, 16–19 (1977).

X-ray bursts and neutron-star thermonuclear flashes

P. C. Joss

Department of Physics and Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Some of the properties of thermonuclear flashes that should occur in the surface layers of accreting neutron stars are investigated. Such flashes may account for many of the observed properties of X-ray burst sources. Helium seems to be the most promising type of nuclear fuel for producing flashes that result in X-ray bursts.

THE history of the discovery of X-ray burst sources has been described by Grindlay¹, and their observed properties have been reviewed by Lewin² and Lewin and Joss³. These sources emit bursts of X-rays with risetimes of $\lesssim 1$ s and decay time scales of ~ 3 –100 s. If they are at a mean distance of ~ 10 kpc, as suggested by their apparent concentration towards the direction of the galactic centre, then the X-ray energy emitted in each burst is typically $\sim 10^{38}$ – 10^{40} erg. The maximum luminosity during the course of a burst is roughly equal to the Eddington limit for a $1 M_{\odot}$ object ($\sim 10^{38}$ erg s⁻¹). The recurrence intervals between bursts are generally in the range $\sim 10^4$ – 10^5 s, except for the 'rapid burster' MXB1730–335 (ref. 4) which has recurrence intervals of ~ 10 –100 s. Many, if not all, burst sources undergo extended inactive states lasting for weeks or longer, during which no bursts are seen. At least eight burst sources also seem to be associated with relatively steady X-ray sources, to within positional accuracies of $\sim 0.1^{\circ}$. The ratio, α , of 'steady' X-ray luminosity to time-averaged burst luminosity is as large as ~ 250 for the associated sources, but an upper limit as small as $\alpha \lesssim 2$ has been placed on one burst source (MXB1730–335) with no identified steady counterpart.

A theoretical model of X-ray burst sources must account for these observations. Some models that have been suggested invoke mechanisms involving massive black holes^{5,6} or instabilities in the accretion on to a compact object of roughly solar mass^{7–12}. Here we discuss another type of model: the production of X-ray bursts by thermonuclear flashes in the freshly accreted matter near the surface of an accreting neutron star.

This model was first proposed by Maraschi and Cavaliere¹³ and Woosley and Taam¹⁴. Even before the discovery of X-ray burst sources, the possibility of thermonuclear flashes on accreting neutron stars was considered by Hansen and Van Horn¹⁵. The present paper investigates further some of the physical processes relevant to such flashes in the context of the rapidly growing volume of observational information. This discussion is still highly preliminary, as realistic numerical calculations of such flashes have not yet been carried out; in particular, the characteristic timescale between flashes is unknown. We conclude that thermonuclear flashes may account for some, but not all, of the observed X-ray burst sources. Our discussion also considers why observational evidence has not yet been obtained for thermonuclear flashes in X-ray pulsars, which are also very likely to be accreting neutron stars. A discussion of some of the same points, including the physics of nuclear fusion processes in accreting neutron stars and the energetics of thermonuclear flashes, has been presented independently by Lamb and Lamb¹⁶.

The physical picture

We consider a neutron star undergoing accretion of mass from a binary stellar companion. The infalling matter will be rich in hydrogen and/or helium. But, in the interior of the neutron star (density $\rho \gtrsim 10^7$ g cm⁻³ or depth $h = R - r \gtrsim 10^2$ cm (see ref. 15),

where r is the distance from the centre of the neutron star and R is the neutron-star radius), nuclear reactions will proceed rapidly even at zero temperature and the nuclei will tend towards a state of nuclear statistical equilibrium. As the infalling matter accumulates, a given element of mass will pass through increasing density and must release the nuclear binding energy associated with the transmutation from its initial composition to a composition of iron-peak elements. If the accreted matter is pure hydrogen, then the energy released will be $\sim 8 \times 10^{18}$ erg g⁻¹.

If the neutron star is cold (internal temperature $< 10^7$ K) and if the gravitational energy released by accretion is radiated away sufficiently rapidly, then the first reaction to proceed will be either pycnonuclear fusion or electron capture by protons at $\rho \simeq 10^7$ g cm⁻³, ultimately producing helium¹⁷. At somewhat higher densities, fusion of helium into heavier elements will proceed. These processes may occur quasi-statically¹⁸, especially if the temperature is very low¹⁷. If, however, either the accretion rate or the internal temperature of the neutron star is sufficiently high, then the fusion will be thermally unstable and should result in a thermonuclear flash; the thermonuclear runaway time scale for such a flash is usually $\lesssim 1$ s (ref. 15). If the energy released in a flash can be transmitted to the neutron-star surface sufficiently rapidly and then lost as electromagnetic radiation, an X-ray burst will result.

We use here the models of Hansen and Van Horn¹⁵ for neutron stars that are undergoing spherically symmetric accretion and steady-state nuclear burning, to estimate the conditions in the surface layers just before a flash. The actual conditions in the surface layers will be somewhat different in the absence of steady-state burning, but more accurate estimates cannot be made until time-dependent calculations of neutron-star envelopes undergoing nuclear burning are carried out.

Energetics

Several useful constraints on thermonuclear-flash models can be derived from simple energetic considerations.

The gravitational energy per unit mass liberated by the accreting matter is given by $E_a = (GM/R) = \eta_a c^2$, where M is the mass of the neutron star, and $\eta_a \sim 0.2$ is the ratio of released gravitational energy to rest-mass energy. (The value of η_a may be as small as $\sim 10^{-3}$ or as large as ~ 0.4 , depending upon the mass of the neutron star and the equation of state of matter at high densities.) This energy will be radiated away from the neutron-star surface with a time-averaged luminosity

$$L_a = \dot{m} E_a \simeq 2 \times 10^{37} \left(\frac{\eta_a}{0.2} \right) \left(\frac{\dot{m}}{10^{17} \text{ g s}^{-1}} \right) \text{ erg s}^{-1} \quad (1)$$

where $\dot{m}$ is the time-averaged mass accretion rate. In the case of the neutron stars that are the X-ray emitting components of binary X-ray pulsar systems, this luminosity is evidently emitted mostly as X radiation.

Nuclear fusion will also contribute to the total luminosity of an accreting neutron star¹⁸. The energy per unit mass released by nuclear fusion can be written as $E_n = \eta_n c^2$, where η_n is the energy efficiency of the reaction; $\eta_n \simeq 7 \times 10^{-3}$ for conventional stellar fusion of pure hydrogen (PP chains or CNO cycle), $\eta_n \simeq 1 \times 10^{-3}$ for helium-burning, and $\eta_n \lesssim 1 \times 10^{-3}$ for the burning of heavier elements. As discussed below, full hydrogen-burning will not occur in a neutron-star thermonuclear flash, and we therefore

expect that in general $\eta_n \lesssim 1 \times 10^{-3}$. If X-ray bursts are caused by thermonuclear flashes in the accreted matter, then the ratio of the time-averaged accretion luminosity to the time-averaged luminosity in bursts is given by

$$\alpha_{\text{flash}} \simeq 2 \times 10^2 \left(\frac{\eta_n}{0.2} \right) \left(\frac{\eta_n}{10^{-3}} \right)^{-1} f^{-1} \xi^{-1} \quad (2)$$

where η_n is the efficiency factor appropriate to the nuclear reactions that participate in the flash, f is the fraction of the nuclear fuel that is consumed in flashes (rather than during the interflash periods), and ξ is the fraction of the flash energy that is transmitted to the neutron-star surface (rather than lost to the interior). The values of f and ξ are unknown, but we shall make the reasonable assumption that both quantities are of order unity. Thus, if the accretion on to the neutron star is steady on a timescale of hours or longer, then we expect that the time-averaged 'steady' X-ray flux from the neutron star will exceed the time-averaged burst flux by at least a factor of $\sim 10^2$.

It is also necessary that the accreted mass be sufficiently large to supply the thermonuclear energy of the flash. If a steady state has been established between the accretion of nuclear fuel by the neutron star and its consumption in flashes, then this constraint can be expressed as

$$\varepsilon_B \simeq 9 \times 10^{38} \left(\frac{\eta_n}{10^{-3}} \right) \left(\frac{\dot{m}}{10^{17} \text{ g s}^{-1}} \right) \left(\frac{\tau}{10^4 \text{ s}} \right) f \xi \text{ erg} \quad (3)$$

where ε_B is the average total energy emitted in each burst and τ is the average time interval between bursts. (Only a fraction of the envelope above the burning shell need be consumed in a given flash, so that the total mass δm of the overlying surface layers may exceed $\dot{m}\tau$ by a substantial factor.) If $\dot{m}\tau$ is actually much less than 10^{21} g , then the flashes will be weak and will not produce observable X-ray bursts. The nuclear energy generation rate required to supply the flash energy ε_B/ξ is given by

$$q_{\text{nuc}} \gtrsim 10^{17} \left(\frac{\varepsilon_B}{10^{39} \text{ erg}} \right) \left(\frac{\delta m}{10^{21} \text{ g}} \right)^{-1} \left(\frac{\tau_B}{10 \text{ s}} \right)^{-1} f^{-1} \xi^{-1} \text{ erg g}^{-1} \text{ s}^{-1} \quad (4)$$

where τ_B is the timescale of duration of the burst emission.

The energy of the flash is thermalised as it is transmitted to the neutron-star photosphere (see below), so that the characteristic photon energy of the burst will be

$$\begin{aligned} e_\gamma &\simeq 4k(\varepsilon_B/4\pi\sigma R^2\tau_B)^{1/4} \\ &\simeq 6 \left(\frac{\eta_n}{10^{-3}} \right)^{1/4} \left(\frac{\dot{m}}{10^{17} \text{ g s}^{-1}} \right)^{1/4} \left(\frac{\tau}{10^4 \text{ s}} \right)^{1/4} \left(\frac{\tau_B}{10 \text{ s}} \right)^{-1/4} \times \\ &\times \left(\frac{R}{10^6 \text{ cm}} \right)^{-1/2} f^{1/4} \xi^{1/4} \text{ keV} \end{aligned} \quad (5)$$

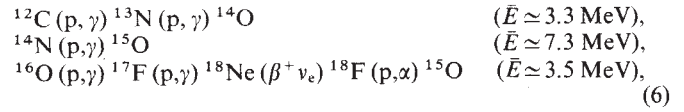
The indicated values of ε_B and e_γ for the chosen plausible parameter values are consistent with the observed properties of X-ray burst sources.

Nuclear reactions

We first consider hydrogen-burning as a possible source of nuclear energy. The density at the base of a steadily burning hydrogen-rich layer will be in the range $10^4 \lesssim \rho_b \lesssim 10^5 \text{ g cm}^{-3}$ (ref. 15). (We shall use ρ_b and T_b to denote the density and temperature at the base of a layer of any specified nuclear fuel.) It seems probable that ρ_b will not lie very far outside this range even if the burning is unsteady.

For the PP chains, the reaction rate is too slow at any temperature to release the energy in a time $\lesssim 10 \text{ s}$, as would be required to account for the short timescales of the observed bursts. The CNO cycle is unable to operate fully, because the β decays of the intermediate nuclei (^{13}N , ^{15}O , and ^{17}F) require timescales of $\sim 10^2$ – 10^3 s . There are no known β -unstable excited states of any of these nuclei that could be thermally populated and thereby appreciably shorten the β -decay timescales, nor will the timescales be effectively reduced, at the expected densities, by captures of continuum electrons. Some of the β -unstable nuclei will undergo

additional proton captures during the flash, provided that the temperatures attained are sufficiently high ($\gtrsim 10^9 \text{ K}$). The resultant reaction chains for initial ^{12}C , ^{14}N , and ^{16}O nuclei are as follows:



where $\bar{E}$ is the average energy liberated per proton capture (excluding the average neutrino loss, if any). The β -decay half life of ^{18}Ne is 1.5 s, so that this nucleus can decay sufficiently rapidly to continue through the reaction chain on a timescale less than τ_B . Of course, ^{13}C , ^{15}N and other heavier nuclei will participate in similar reaction chains, but their initial abundances are likely to be lower than those of ^{12}C , ^{14}N , and ^{16}O . (The daughter nuclei of ^{14}O and ^{15}O may be important if the same layer of matter participates in more than one flash, but this consideration does not substantially alter the expected values of η_n .)

An average of ~ 2 protons will be consumed for each CNO nucleus, and the amount of energy liberated per CNO nucleus will be $\sim 8 \text{ MeV}$. Hence, if the duration of flash-burning is no greater than the observed duration of an X-ray burst, then the energy efficiency of the protonisation of CNO nuclei will be $\eta_n \sim 9 \times 10^{-3} N_{\text{CNO}}$, where N_{CNO} is the fractional abundance by number of CNO nuclei in the unburned matter. (We assume that $N_{\text{CNO}} \lesssim 0.3$.) This value of η_n will be increased slightly by proton captures on to other heavy nuclei and will also be increased by α -captures, which should occur at the high temperatures attained in the flash. The value of η_n will thus depend on the temperatures achieved during the flash, as well as on whether the same layer of accreted matter participates in more than one flash. If the CNO abundance is near the cosmic value of $N_{\text{CNO}} \sim 10^{-3}$, however, then the energy efficiency of a hydrogen-burning flash should generally be quite low.

If, instead, the flash is caused by the fusion of helium or heavier elements, then the reaction rates will not be retarded by the effect of weak interactions and the reactions can proceed very swiftly ($\ll 1 \text{ s}$) when sufficiently high temperatures are achieved. (The relevant nuclear reaction rates are not greatly affected by relativistic effects, electron degeneracy effects, or electron screening, so that conventional reaction-rate formulae may be used.) The energy efficiency of helium-burning is no lower than that of hydrogen-burning, unless the CNO abundance in a hydrogen-burning flash is extremely high ($N_{\text{CNO}} \gtrsim 0.1$).

The value of η_n in a flash of helium or heavier elements may be increased by convective mixture. For example, such mixing could draw hydrogen into a helium-burning shell and result in proton-captures on to the heavier nuclei that are being synthesised.

Neutrino loss rates will be highest if the flash occurs in helium. The models of ref. 15 suggest that $\rho_b \gtrsim 10^6 \text{ g cm}^{-3}$ at the base of the helium-rich layer; the released energy would be sufficient to raise the temperature to $T_b \sim 2 \times 10^9 \text{ K}$ for $\rho_b \sim 10^6 \text{ g cm}^{-3}$ (where a substantial fraction of the internal energy would be taken up by blackbody radiation) and to $T_b \sim 10^{10} \text{ K}$ for $\rho_b > 10^7 \text{ g cm}^{-3}$ (where the radiation energy density would be negligible). The neutrino loss rates (ref. 19; ref. 20 and references therein) are as high as $\sim 10^{18} \text{ erg g}^{-1} \text{ s}^{-1}$ at $\rho_b \sim 10^7 \text{ g cm}^{-3}$ and $T_b \sim 10^{10} \text{ K}$, but decrease rapidly with decreasing temperature and also decrease with increasing density. The effective helium-burning flash energy and the value of T_b may thus be slightly reduced in a narrow range of densities near $\rho_b = 10^7 \text{ g cm}^{-3}$, but will be unaffected at other densities. At the higher densities anticipated for a flash of heavier elements ($\rho_b \gtrsim 10^9 \text{ g cm}^{-3}$ and $T_b \sim 10^{10} \text{ K}$; see ref. 14), the neutrino loss rates are negligible. They will also be negligible in a hydrogen-burning flash, since T_b will then be no greater than $\sim 10^9 \text{ K}$ unless N_{CNO} is extremely large ($\gtrsim 0.1$).

Heat transport

If a thermonuclear flash is to produce an observable X-ray burst, then the surface layers of the neutron star must be sufficiently thick

to provide the nuclear energy of the burst. One must, therefore, also consider whether the timescale for transport of the energy of the flash to the neutron-star surface will be $\lesssim 1$ s, as required to account for the observed rapid risetimes of the X-ray bursts.

Conduction of heat by electrons is probably unimportant even if the electron degeneracy of the surface layers is high before the flash, because the degeneracy will generally be lifted and the conductivity will become small in each layer of matter before most of the heating of the layer by the flash has occurred. Convection should be important in the development of the flash, and an estimate based on the mixing-length theory indicates that the timescale for the initial transport of heat through the convectively unstable region will not greatly exceed the sound travel time ($\ll 1$ s). In many possible circumstances, however, convection will not transport the energy of the flash all the way to the neutron-star photosphere. We demonstrate the validity of this statement as follows.

Consider first a hydrogen-burning flash. If all of the energy is released rapidly (in a time $\ll 1$ s), and the CNO abundance is enhanced (say, $N_{\text{CNO}} \sim 2 \times 10^{-2}$), then ~ 4 MeV is released for each proton consumed, and a maximum temperature of $T_b \sim 10^9$ K may be achieved at the base of the burning shell. This result is self-consistent, since the proton-capture reactions will, in fact, proceed very rapidly at $T_b \sim 10^9$ K. Some burning of helium and heavier elements should also occur, but the temperature will not be increased greatly over this value. At this temperature, electron degeneracy will be lifted and the flashed matter will behave as a mixture of a nearly ideal gas plus blackbody radiation. We distinguish between two cases. (1) If $\rho_b > 10^4$ g cm $^{-3}$, then the internal energy of the gas dominates that of the radiation and

$$\beta \equiv \frac{\text{gas pressure}}{\text{total pressure}} \gtrsim 0.9$$

Hence, as a convective element rises and cools nearly adiabatically, its temperature, T_{conv} , will depend on its density, ρ_{conv} , as

$$T_{\text{conv}} \propto \rho_{\text{conv}}^{2/3} \quad (7)$$

(2) If, instead, $\rho_b < 10^3$ g cm $^{-3}$, then the radiation energy density becomes significant, so that $\beta \lesssim 0.3$, T_b is somewhat reduced, and

$$T_{\text{conv}} \propto \rho_{\text{conv}}^{1/3} \quad (8)$$

For the greater energy per unit mass that might be released in a flash of helium or heavier elements, degeneracy is again generally lifted; for $\rho_b > 10^8$ g cm $^{-3}$, one obtains $\beta \gtrsim 0.9$ and $T_b \sim 10^{10}$ K, while for $\rho_b < 10^7$ g cm $^{-3}$, one obtains $\beta \lesssim 0.3$ and somewhat smaller values of T_b .

In all cases where $\beta \gtrsim 0.9$, relation (7) implies that the temperature of a convective element will have fallen to $T_p \lesssim 10^6$ K due to its expansion by the time it has reached the photosphere. This may well be less than the ambient photospheric temperature T_e , in which case the buoyancy of a convective element will vanish at a deeper level of the surface layers and no convective heat flux will reach the photosphere. Even if T_p is in excess of T_e , it is smaller by a factor of $\gtrsim 10$ than the photospheric temperature required to produce the X-radiation of a burst (see equation (5)). On the other hand, if $\beta \lesssim 0.3$, then relation (8) implies that $T_p \gtrsim 10^7$ K, so that convection can transport heat to the photosphere with adequate efficiency to produce an X-ray burst. But, for a flash of elements heavier than helium, β should always be close to unity and convection should, therefore, never be able to transport heat all the way to the photosphere (compare ref. 14).

The remaining heat transport mechanism of significance is radiative diffusion. From the equation of radiative transfer in local thermodynamic equilibrium, we find that under the conditions of a flash the radiative luminosity, L_{rad} , exceeds the Eddington limit L_{crit} . This implies that the heat transport will be dominated by convection up to a level where the temperature gradient is sufficiently small to allow $L_{\text{rad}} \approx (1 - \beta)L_{\text{crit}}$ (ref. 21). If $\beta \lesssim 0.3$ at

the burning shell and convection therefore penetrates to the photosphere, the radiative luminosity escaping from the photosphere should still be $\sim L_{\text{crit}}$. For the outer portion of the surface layers, we may thus take

$$L_{\text{rad}} \sim (1 - \beta)L_{\text{crit}} \\ \approx 1 \times 10^{38} (1 - \beta) \left(\frac{M}{M_{\odot}} \right) \left(\frac{\kappa}{0.4 \text{ cm}^2 \text{ g}^{-1}} \right)^{-1} \text{ erg s}^{-1} \quad (9)$$

where β is the typical value of β and κ is the typical radiative opacity in the outer surface layers (expressed in units of the Thomson scattering opacity for pure hydrogen).

Let ρ_r be the density at the base of the radiative layer when the convective motions first penetrate to that layer. Using equation (7) and the equations of hydrostatic equilibrium and radiative transfer, and assuming that L_{rad} and κ vary slowly through the outer surface layers before the flash, we can estimate ρ_r for cases where $\beta \gtrsim 0.9$ at the burning shell:

$$\rho_r \sim 1 \times 10^5 \left(\frac{\rho_b}{10^8 \text{ g cm}^{-3}} \right)^2 \left(\frac{T_b}{10^{10} \text{ K}} \right)^{-3} \left(\frac{T_o}{10^7 \text{ K}} \right)^3 \text{ g cm}^{-3} \quad (10)$$

Here, T_o is the temperature before the flash at the level where $\rho = 10^2$ g cm $^{-3}$. The indicated values of ρ_b and T_b are plausible for a helium-burning flash¹⁵, and we expect that $T_o \sim 10^7$ K in a neutron star undergoing a moderate rate of accretion. Using equation (10) and the equation of hydrostatic equilibrium, we further find that the timescale for the initial diffusion of radiation from the base of the radiative layer to the photosphere is

$$\tau_r \lesssim \kappa \rho_r h_r^2 / c \\ \sim 0.1 \left(\frac{\kappa}{0.4 \text{ cm}^2 \text{ g}^{-1}} \right) \left(\frac{\rho_b}{10^8 \text{ g cm}^{-3}} \right)^{10/3} \left(\frac{T_b}{10^{10} \text{ K}} \right)^{-5} \left(\frac{T_o}{10^7 \text{ K}} \right)^7 \\ \times \left(\frac{M}{M_{\odot}} \right)^{-2} \left(\frac{R}{10^6 \text{ cm}} \right)^4 \mu_r^{-2} \text{ s} \quad (11)$$

Here h_r is the depth of the initially radiative layer and μ_r is the typical mean molecular weight in the layer.

These estimates of L_{rad} and τ_r are crude, but they indicate that convection and/or radiative diffusion should transport sufficient heat from a helium-burning flash to the neutron-star photosphere to produce an X-ray burst and that the initial risetime of the burst should be sufficiently short to agree with the observations. From equation (3), we also find that the indicated value of L_{rad} is just sufficient to give a timescale for burst emission of $\tau_B \sim \epsilon_B / L_{\text{rad}} \sim 10$ s. The luminosity and diffusion timescale for a hydrogen-burning flash are also capable of accounting for the properties of an X-ray burst, but ϵ_B and τ_B will be too small and α_{flash} will be too large unless the CNO abundance is considerably enhanced over its cosmic value. A thermonuclear flash of carbon or heavier elements at $\rho_b \gtrsim 10^9$ g cm $^{-3}$, $T_b \sim 10^{10}$ K, and depth $h_b \sim 10^4$ cm (ref. 14) can reproduce the emitted energies of X-ray bursts, but may not be able to produce sufficiently short burst risetimes.

Convection will carry fresh fuel into the burning shell to feed the flash. The total amount of available fuel may therefore exceed the mass of the initial burning shell by a large factor. On the other hand, the burning will tend to be terminated by the expansion and cooling of the surface layers following the initial flash. In the absence of detailed numerical calculations, one cannot determine what the duration of the flash will be. It is plausible, however, that the expansion of the surface layers and subsequent termination of the flash will occur on a thermal timescale for the surface layers:

$$\tau_{\text{th}} \sim 10 \left(\frac{2 - \beta_s}{\beta_s} \right) \left(\frac{T_s}{10^{10} \text{ K}} \right) \left(\frac{\delta m}{10^{21} \text{ g}} \right) \left(\frac{L_{\text{rad}}}{10^{38} \text{ erg s}^{-1}} \right)^{-1} \mu_s^{-1} \text{ s} \quad (12)$$

where β_s , T_s , and μ_s are the typical values of β , temperature, and

mean molecular weight in the surface layers following the flash. (Note that τ_{th} must be comparable to τ_{B} , since the internal energy of the surface layers is dominated by the energy released in the flash.) If all of the available nuclear fuel is consumed in a time less than τ_{th} , then the flash will instead terminate on a nuclear-exhaustion timescale:

$$\tau_{\text{nuc}} \sim 10 \left(\frac{\eta_{\text{n}}}{10^{-3}} \right) \left(\frac{\delta m}{10^{21} \text{ g}} \right) \left(\frac{L_{\text{nuc}}}{10^{38} \text{ erg s}^{-1}} \right)^{-1} f \text{ s} \quad (13)$$

where L_{nuc} is the integrated nuclear energy generation rate throughout the surface layers; one expects $L_{\text{nuc}} > L_{\text{rad}}$ during the flash. Either (or both) of these timescales can be comparable to or shorter than 10 s, as desired if thermonuclear flashes are to account for the observed bursts.

The emergent radiation will be effectively thermalised. A ~ 10 MeV photon produced in a (p, γ) reaction, for example, will be thermalised to ~ 10 KeV after $\sim 10^2$ scattering events. Hence, most photons generated at an optical depth greater than ~ 10 will be thermalised by the time they reach the photosphere. The presence of convection should not inhibit the thermalisation; from the mixing-length theory, we estimate that a photon produced in the burning shell will generally undergo many scattering events before reaching the photosphere. Moreover, even if thermonuclear fusion spreads through most of the surface layers during the progress of the flash, a significant amount of fusion should not reach an optical depth as small as ~ 10 . The transport of radiation through the accreting matter, however, may complicate the emergent spectrum, as seems to be the case for the binary X-ray pulsars (see ref. 22 and references therein).

Conclusions

The surface layers of accreting neutron stars should undergo thermonuclear flashes that will produce unsteady radiation, and this radiation may have properties similar to those of the observed X-ray burst sources. Helium-burning flashes are the most promising type for producing X-ray bursts, because of the relatively high energy yield and because the timescale for the released heat to reach the neutron-star photosphere is shorter than for flashes that result from the burning of carbon and other heavier elements. Enhanced CNO abundances ($N_{\text{CNO}} \geq 10^{-2}$) are needed if hydrogen-burning flashes are to be sufficiently energetic.

The burning of heavier elements within an accreting neutron star is likely to be thermally unstable under a wide range of conditions¹⁴. The effects of such instabilities could be more difficult to observe than the X-ray bursts detected, since the radiation may well be emitted at characteristic photon energies considerably less than 10 KeV and on timescales considerably longer than 10 s.

For thermonuclear flashes under nearly all possible circumstances, the value of α_{flash} will be ≥ 10 ; the most plausible values of α_{flash} are $\geq 10^2$. (The smallest values of α_{flash} will be attained for low-mass neutron stars, with small values of η_{n} , or for hydrogen-burning flashes involving highly enhanced CNO abundances, with large values of η_{n} .) Hence, a thermonuclear-flash model cannot account for MXB1730-335 ($\alpha \leq 2$) unless one inserts an additional *ad hoc* complication into the model, such as highly anisotropic beaming of the steady emission so that it largely misses the Earth. But, the empirical difference between the behaviour of MXB1730-335 and other burst sources (refs 2, 3 and references therein) suggests that the fundamental nature of this source may be unique. The burst source MXB1659-29, which has recently been found to have $\alpha < 25$ (W. Lewin, J. Hoffman, and J. Doty, unpublished SAS-3 data), may also pose a serious problem for thermonuclear-flash models.

Another difficulty with thermonuclear-flash models is their prediction of simple blackbody spectra for X-ray bursts, in contrast to the rather complex changes in spectrum that are displayed during the progress of some observed bursts^{2,3}. To account for such behaviour, it is necessary to invoke time-dependent reprocessing of the burst radiation by matter above the

Alfvén surface of the neutron star. It is worth noting, however, that the time-averaged burst spectra more closely resemble a blackbody than a thin-bremsstrahlung or power-law spectrum². Moreover, the instantaneous spectra of at least three burst sources seem to be well-represented by a blackbody with an effective radius of ~ 10 km throughout the course of each burst²³⁻²⁵.

On the other hand, thermonuclear-flash models may provide a natural explanation for the quasi-periodic character of the recurrence intervals that have been observed in several burst sources^{2,3}. If the accretion rate on to the neutron star is nearly constant on a timescale of days and each burst consumes about the same amount of fuel, then a nearly constant interval between bursts may result (since a specific envelope mass may be needed to precipitate a flash). The observed variations in the recurrence intervals on timescales of several days to weeks could result from variations in the accretion rate on such timescales, while the cessation of bursting in many sources for periods of months or longer may be caused by either (1) a sharp reduction in the accretion rate or (2) an increase in the accretion rate that changes the properties of the neutron-star surface layers so as to alter the character of the nuclear fusion processes.

It is somewhat puzzling that all accreting neutron stars do not display some evidence of thermonuclear flashes. There are 12 known X-ray pulsars, which are widely believed to be accreting neutron stars (see ref. 22 and references therein) but which have not been observed to burst. It is possible that the X-ray burst sources are intrinsically very similar to the X-ray pulsars, but with lower accretion rates or other differences in system parameters¹⁰. But, in all cases investigated by Hansen and Van Horn¹⁵, the nuclear burning shells were either thermally unstable or at most marginally stable. These models do show a slight trend toward greater stability of nuclear burning as the mass accretion rate increases, but instability was found even at accretion rates corresponding to steady X-ray luminosities of as much as $\sim 10^4 L_{\odot}$. A better understanding of these problems must await a detailed time-dependent numerical treatment of nuclear fusion processes in neutron-star surface layers.

As noted above, it is possible that more than one mechanism is responsible for the observed burst sources. In order to distinguish between models invoking massive black holes^{1,5,6} and those which involve moderate-mass compact objects⁷⁻¹² that may be in binary systems, observations of X-ray pulsing, X-ray eclipses, or optical companion stars which exhibit binary characteristics (ellipsoidal light variations or orbital velocity curves) could prove decisive.

It is equally important and, in principle, more straightforward to distinguish between thermonuclear-flash models for burst sources and those models which involve instabilities in accretion on to a neutron star⁷⁻¹¹. This could be accomplished by determining the distribution of values of α for a large statistical sample of sources. If the value of α for most sources were $\geq 10^2$, strong support for thermonuclear-flash models would be provided. If, on the other hand, there exists a relatively continuous distribution of values of α extending well below 10^2 , then one could reasonably conclude that nuclear flash-burning is not responsible for any of the burst sources.

I thank many colleagues, especially Saul Rappaport, for helpful discussions and Professor Martin Rees and the Institute of Astronomy for their hospitality during August 1977. This work was supported in part by NASA (contract NAS5-11450). The author is an Alfred P. Sloan Research Fellow.

Received 20 August; accepted 2 September 1977.

- Grindlay, J. E. *Comm. Astrophys. Space Phys.* **6**, 165 (1977).
- Lewin, W. H. G. *Mon. Not. R. Astr. Soc.* **179**, 43 (1977); *Ann. N.Y. Acad. Sci.* (in the press).
- Lewin, W. H. G. & Joss, P. C. *Nature* **270**, 211-216 (1977).
- Lewin, W. H. G. *et al. Astrophys. J. Lett.* **207**, L95 (1976).
- Grindlay, J. & Gursky, H. *Astrophys. J. Lett.* **205**, L131 (1976).
- Bahcall, J. N. & Ostriker, J. P. *Nature* **262**, 37 (1976).
- Svestka, J. *Astrophys. Space Sci.* **45**, 21 (1976).
- Henriksen, R. N. *Astrophys. J. Lett.* **210**, L19 (1976).
- Baan, W. A. *Astrophys. J.* **214**, 245 (1977).
- Joss, P. C. & Rappaport, S. *Nature* **265**, 222 (1977).
- Lamb, F. K., Fabian, A. C., Pringle, J. E. & Lamb, D. Q. *Astrophys. J.* (in the press).
- Wheeler, J. C. *Astrophys. J.* **214**, 560 (1977).
- Maraschi, L. & Cavaliere, A. *Highlights of Astronomy* **4** (in the press).
- Woosley, S. E. & Taam, R. E. *Nature* **263**, 101 (1976).
- Hansen, C. J. & Van Horn, H. M. *Astrophys. J.* **195**, 735 (1975).

16. Lamb, D. Q. & Lamb, F. K. *Astrophys. J.* (submitted).
17. Van Horn, H. M. & Hansen, C. J. *Astrophys. J.* **191**, 479 (1974).
18. Rosenbluth, M. N. *et al. Astrophys. J.* **184**, 907 (1973).
19. Beaudet, G., Petrosian, V. & Salpeter, E. F. *Astrophys. J.* **150**, 979 (1967).
20. Barkat, Z. A. *Rev. Astr. Astrophys.* **13**, 45 (1975).
21. Joss, P. C., Salpeter, E. E. & Östriker, J. P. *Astrophys. J.* **181**, 429 (1973).

22. Rappaport, S. & Joss, P. C. *Nature* **266**, 123 (1977).
23. Swank, J. H. *et al. Astrophys. J. Lett.* **212**, L73 (1977).
24. Hoffman, J., Lewin, W. H. G. & Doty, J. *Mon. Not. R. astr. Soc.* **179**, 57P (1977).
25. Hoffman, J. A., Lewin, W. H. G., Doty, J., Jernigan, J. G. & Haney, M. *Astrophys. J. Lett.* (in the press).

Ovalbumin gene is split in chicken DNA

R. Breathnach, J. L. Mandel & P. Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 44 de l'INSERM et Institut de Chimie Biologique, Faculté de Médecine, Strasbourg, France

The ovalbumin gene is split in chicken DNA. Two interruptions in the sequences coding for ovalbumin mRNA have been detected, at least one of them lying in the protein coding sequence. The unexpected gene organisation is present both in oviduct cells highly specialised in ovalbumin synthesis and in erythrocytes.

DNA SEQUENCES coding for the variable and constant regions of immunoglobulin molecules are widely separated in embryo cells but are much closer in differentiated lymphocytes¹. This finding suggests that gene rearrangement could be one of the mechanisms involved in cell differentiation. Chicken oviduct is particularly useful for the study of such a possibility, since administration of oestrogen to immature chicks results in cytodifferentiation of tubular gland cells, which comprise up to 90% of the cells in the magnum portion of laying hen oviduct and which synthesise the major egg white proteins including ovalbumin, ovomucoid, conalbumin and lysozyme (see refs 2, 3). In particular, ovalbumin accounts for 50–65% of total protein synthesis in tubular gland cells, and the rate of transcription of its mRNA (which represents 40% of the total mRNA population) is under hormonal control^{3,4}.

As reported for integrated viral genes⁵, it is possible to construct a physical map of restriction endonuclease sites within and around a single-copy gene. Briefly, this involves cleavage of the DNA with various restriction endonucleases, fractionation of the fragments by electrophoresis, their transfer to nitrocellulose sheets⁶, hybridisation to a specific probe labelled *in vitro* by nick-translation⁷ to high specific activity and location of the gene fragments by autoradiography. In this study we have used the bacterial plasmid, pCR1ov2.1, which contains about 80% of the nucleotide sequences present in ovalbumin mRNA⁸ as a hybridisation probe to investigate the possibility that the ovalbumin gene environment could be different in DNA prepared from laying hen oviduct or erythrocytes. Unexpectedly, we have found that the DNA sequences complementary to ovalbumin mRNA are split into several fragments in oviduct DNA, and that the same peculiar ovalbumin gene organisation is present in laying hen oviduct and erythrocyte DNA.

Mapping chicken ovalbumin gene in oviduct and erythrocyte DNA

The *HhaI* fragment A (termed *Hhaov*) of the ovalbumin clone pCR1ov2.1 is about 2,400 base pairs long and contains a sequence of about 1,730 base pairs coding for ovalbumin mRNA⁸. After nick-translation with ³²P-labelled deoxyribonucleoside triphosphates⁷, this fragment serves as a filter hybridisation probe to detect DNA fragments containing the ovalbumin gene in restriction endonuclease digests of chicken DNA, using the blotting technique developed by Southern⁶.

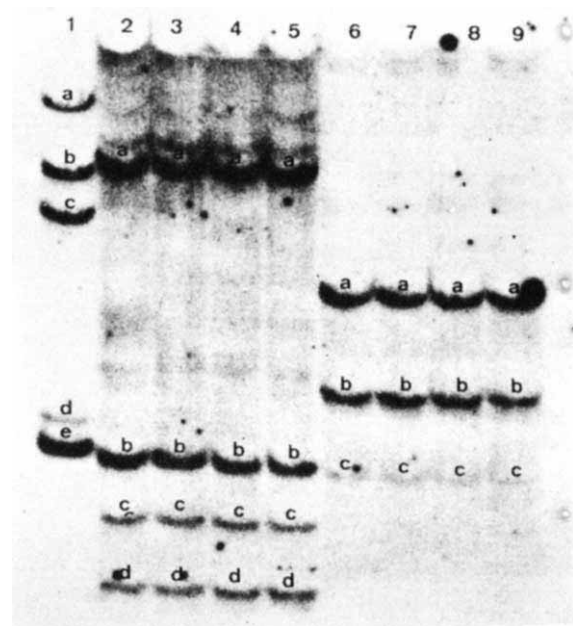


Fig. 1 Detection of ovalbumin gene sequences in chicken DNA cleaved with *EcoRI* or *HindIII*. DNA from the magnum portion of laying hen oviduct or from erythrocyte nuclei was purified as in refs 10 and 11, respectively. DNA samples were digested to completion with the restriction enzymes *EcoRI* or *HindIII* (gifts of Dr J. Sümegi). To monitor the extent of digestion, phage λ DNA was included in the incubations. Some digests were treated with proteinase K, phenol extracted, ethanol precipitated and redigested in the presence of fresh phage λ DNA to ensure complete cleavage. Samples of the digested DNA (10 μ g) were electrophoresed on 0.7% Agarose gels (17 cm long, 4 mm thick, lane width 8 mm) at 2 V cm^{-1} , transferred to nitrocellulose filters (Schleicher and Schüll BA85) by a modification of the method of Southern⁶, and fixed to the filter by baking in a vacuum at 80 °C for 3 h. Before hybridisation, filters were soaked at 68 °C in 5 $\times$ SSC, 0.1 M sodium phosphate pH 7.0 containing 0.02% each of Ficoll, bovine serum albumin and polyvinylpyrrolidone¹². Hybridisation to ³²P-labelled *Hhaov* fragment (0.1 μ g, 0.5–2 $\times$ 10⁷ c.p.m.) labelled by nick-translation¹³ was carried out in the soaking solution for 16 h at 68 °C (ref. 5). The filters were washed first with the soaking solution containing 0.5% SDS, then extensively with 1 $\times$ SSC, 0.1 M sodium phosphate pH 8.4, containing 0.5% SDS. All washing was done at 68 °C. Autoradiography was for 7 d at room temperature with a Kodirex film (Kodak). *Hhaov* fragment was purified from an *HhaI* digest of pCR1 ov2.1 DNA by sedimentation in a 5–20% sucrose gradient. Lane 1, Internal markers for molecular weight and hybridisation were prepared from pCR1ov2.1 DNA by cleavage with various restriction nucleases. Only fragments containing ovalbumin gene sequences are revealed after hybridisation. The molecular weight of the fragments was determined using adenovirus 2 DNA cleaved by *EcoRI* or *BamHI* as standard¹⁴. The enzymes generating the observed bands and the sizes of the fragments (in kilo base pairs) are: a, *HindIII*, 13.5; b, *SmaI*, 8.3; c, *HindIII/PstI*, 6.5 (contains the sequences present in *PstA* fragment (see Fig. 2)); d, *HindIII/PstI*, 2.6 (contains sequences present in *PstB* fragment (see Fig. 2)); e, *HhaI*, 2.4. Lanes 2–5, *EcoRI* digests of erythrocyte (lanes 2, 3) or oviduct (lanes 4, 5) DNA. Samples in lanes 3 and 5 were digested twice (see above). The sizes of the fragments hybridising to *Hhaov* are (in kilo base pairs): a, 9.5; b, 2.35; c, 1.8; d, 1.3. Lanes 6–9, *HindIII* digest of erythrocyte (lanes 6, 7) or oviduct (lanes 8, 9) DNA. Samples in lanes 7 and 9 were digested twice (see above). The sizes of the fragments hybridising to *Hhaov* are (in kilo base pairs): a, 4.9; b, 3.2; c, 2.3.

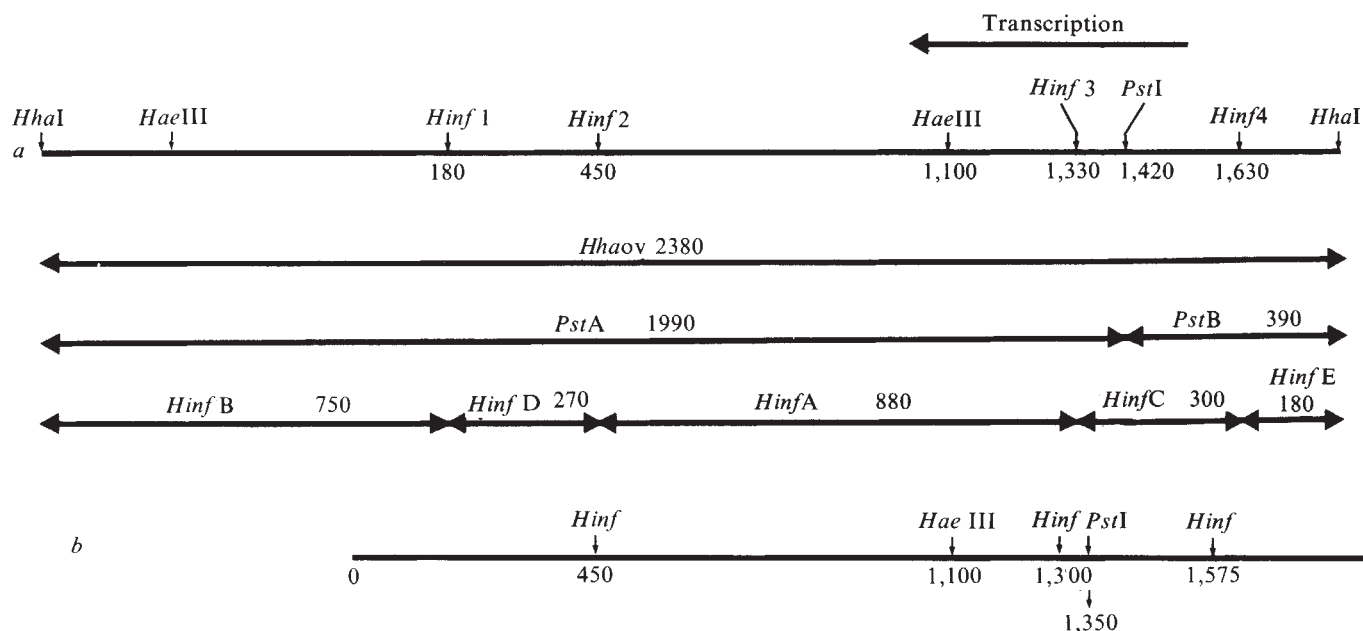


Fig. 2 *a*, Physical map of restriction enzyme sites in the *Hhaov* fragment and orientation of the ovalbumin double-stranded cDNA insert (J.L.M., in preparation). Only the sites relevant to the present study are shown (for convenience, restriction endonuclease *HinfI* is termed *Hinf*). The position of the restriction endonuclease sites on the *Hhaov* fragment is given from the 3' end of the mRNA (excluding the poly (A) tail) assuming that the site *HinfI* corresponds to that predicted from the known sequence of ovalbumin mRNA around position 180 from its 3' end¹⁵. This assumption is supported by the demonstration that the *HinfB* fragment hybridises to mRNA (J.L.M., in preparation). Since the cloned ovalbumin ds cDNA sequence is about 1,730 nucleotides long⁸, all restriction sites from *HinfI* to *Hinf4* inclusive are located in a DNA sequence complementary to ovalbumin mRNA. Location and sizes of all fragments that have been used as hybridisation probes are also shown. *b*, Physical map of relevant restriction endonuclease sites of *in vitro* synthesised ovalbumin ds cDNA as established by Monahan *et al.*⁹. The position of the restriction enzyme sites is given from the 3' end of the mRNA. The two maps are in very good agreement when taking into account that different electrophoresis gel systems and molecular weight markers were used in the two studies.

High molecular weight DNA was prepared from chicken oviduct and erythrocytes, digested with the restriction enzymes *EcoRI* or *HindIII*, electrophoresed on Agarose gels, denatured, and transferred to nitrocellulose filters⁶. After hybridisation with denatured ³²P-labelled *Hhaov* fragment and autoradiography, several bands were seen as shown in Fig. 1, suggesting that sequences complementary to ovalbumin mRNA are present in different DNA fragments. The autoradiography band pattern was identical for oviduct and erythrocyte DNA (Fig. 1) and also for 5-d-old chick embryo DNA (not shown). This multiplicity of hybridisation bands was unexpected, since there is no *EcoRI* or *HindIII* restriction enzyme site in the cloned ovalbumin cDNA sequences (Fig. 2). Absence of such sites in *in vitro* synthesised ovalbumin double-stranded cDNA (ds cDNA) was also recently reported by Monahan *et al.*⁹. The very good agreement between the restriction maps of the cloned and *in vitro* synthesised ds cDNAs (Fig. 2), together with our previous results⁸, exclude the possibility that the multiplicity of the bands can be attributed to DNA sequence rearrangements which might have occurred during the cloning and/or amplification of the ds cDNA. The following results exclude the possibility that the multiple hybridisation bands could be related to incomplete restriction enzyme digestion of the chicken DNA or to nonspecific hybridisation: (1) there was no change in the pattern when the DNA was digested twice, while the phage λ DNA internal marker was digested each time to completion (Fig. 1); (2) the hybrid bands were present even in much more stringent washing conditions (0.1 SSC, 0.5% sodium dodecylsulphate, 68 °C) and were as stable as the hybrids formed with the internal hybridisation markers (see legend to Fig. 1); (3) no hybridisation bands were observed when *EcoRI* DNA fragments of high molecular weight calf thymus DNA were used instead of *EcoRI* DNA fragments of high molecular weight calf thymus DNA were used instead of *EcoRI* fragments of chicken DNA (not shown); (4) the same bands were observed when the *Hhaov* probe was replaced by a probe from which plasmid sequences and GC tails adjacent to the ovalbumin sequences were removed by hy-

bridisation with ovalbumin mRNA followed by S1 nuclease treatment, as described in the legend to Fig. 4.

Hybridisation to probes complementary to defined parts of ovalbumin mRNA

To investigate the significance of the multiple hybridisation bands we have prepared the ³²P-labelled fragments *PstA*, *PstB* and *HinfA* (see legend to Fig. 3). As illustrated in Fig. 2, the *PstA* probe contains most of the DNA sequences complementary to the first 1,420 nucleotides of the mRNA, starting from its 3' end. The *PstB* probe contains DNA sequences complementary to the 5' region of the mRNA from positions 1,420 to at least 1,630 from the 3' end of the mRNA. The *HinfA* probe contains only sequences complementary to the central part of the mRNA (Fig. 2). The hybridisation patterns of ³²P-labelled *PstA* and *PstB* probes to *EcoRI* or *HindIII* restriction fragments of oviduct DNA are shown in Fig. 3. The *PstA* probe hybridised with *EcoRI* fragments a, c and d, but not with fragment b (lane 3). The *HinfA* probe gave results very similar to those obtained with the *PstA* probe (not shown). On the other hand, the *PstB* probe hybridised strongly to the *EcoRI* fragment b, and only weakly to the *EcoRI* fragment a (lane 7). A similar pattern was found with *HindIII* fragments, since the *HindIII* fragment b did not hybridise with *PstA* (lane 4) or *HinfA* (not shown) probes, whereas it was responsible for the main band with the *PstB* probe (lane 8); this latter probe hybridised only weakly to *HindIII* fragment a. But, the *PstB* probe seemed to be contaminated with *PstA* sequences. Annealing of the *PstB* probe to the pCR1ov2.1 fragments serving as internal markers showed indeed that it hybridised with both fragments c and d (Fig. 3, lane 5), which contain the *PstA* and the *PstB* DNA sequences respectively (J. L. Mandel, in preparation and legend to Fig. 1). On the other hand, the *PstA* probe hybridised only to the internal marker fragment c and not to the

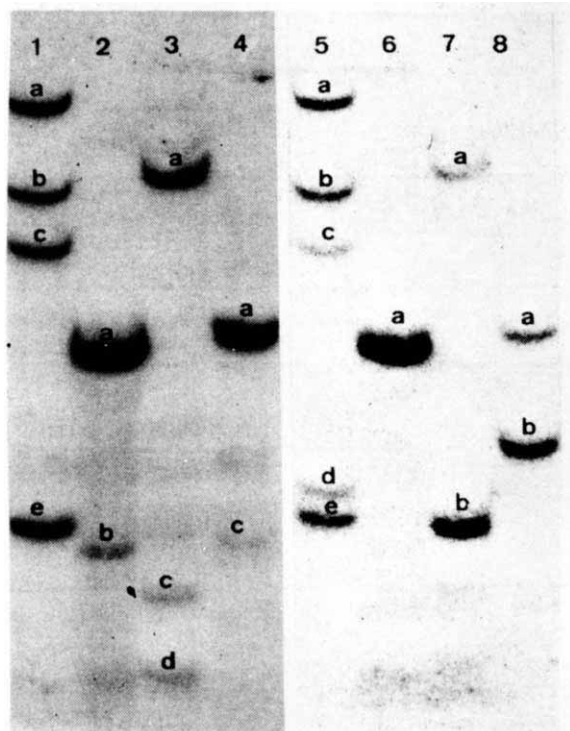


Fig. 3 Detection of fragments containing sequences complementary to defined parts of ovalbumin mRNA in digests of chicken oviduct DNA with *Eco*RI, *Hind*III or *Pst*I. DNA samples were prepared as described in Fig. 1, except that SV40 DNA was used to monitor the digestions with *Pst*I (Biolabs). *Pst*A and *Pst*B probes were labelled by nick-translation of a *Pst*I digest of *Hha*ov. After electrophoresis on a 2% Agarose gel, the labelled *Pst*A and *Pst*B fragments were extracted from the gel¹⁶. Their specific activity was about 1×10^8 c.p.m. μg^{-1} . Autoradiography was carried out with preflashed Kodak RP Royal X-Omat film and intensifying screen at -70°C for 7 d. All other procedures were as described in Fig. 1. Lanes 1-4, hybridisation to ^{32}P -*Pst*A. Lanes 5-8, hybridisation to *Pst*B; lanes 1 and 5, internal markers (see Fig. 1). Lanes 2 and 6, *Pst*I digest of chicken DNA—sizes of the fragments in kilo base pairs; a, 4.7; b, 2.2. Lanes 3 and 7: *Eco*RI digest of chicken DNA (see Fig. 1 for fragment sizes). Lanes 4 and 8, *Hind*III digest of chicken DNA (see Fig. 1 for fragment sizes).

marker fragment d (Fig. 3, lane 1). This indicates that the *Pst*A probe did not contain sequences complementary to *Pst*B, and that the hybridisation of the *Pst*B probe to the marker fragment c was due to a contamination of that probe. This contamination could be due to some degradation of the *Pst*A fragment during ^{32}P -labelling, since the two probes were separated by gel electrophoresis after the nick translation reaction (see legend to Fig. 3). In any case, assuming that the weak hybridisation of the *Pst*B probe to the *Eco*RI and *Hind*III fragments 'a' were due to such a contamination, our results with probes *Pst*A and *Pst*B suggested the possibility that, in the cellular DNA, some sequences complementary to the 5' moiety of the ovalbumin mRNA are physically separated from those complementary to the rest of the mRNA.

Confirmation that the chicken ovalbumin gene is split

DNA fragments from *Eco*RI/*Hind*III, *Eco*RI/*Pst*I and *Hind*III/*Pst*I double digests were hybridised to ^{32}P -labelled probes complementary to different regions of the ovalbumin mRNA (Fig. 4). As expected, multiple hybridisation bands were revealed by autoradiography. Their pattern was identical for oviduct and erythrocyte DNA (data not shown). Irrespective of the combination of restriction endonucleases, there was always one DNA fragment which hybridised strongly to *Pst*B, but not to

*Pst*A or *Hinf*A: band 'b' for the *Eco*RI/*Hind*III digest (compare lanes 1, 4, 6, 8) and for the *Hind*III/*Pst*I digest (lane 3, hybridisation with *Pst*A, *Pst*B and *Hinf*A probes are not shown), band 'c' for the *Eco*RI/*Pst*I digest (compare lanes 2, 5, 7 and 9). From the hybridisation pattern with *Pst*A and *Hinf*A, it is clear that fragments 'a' of *Eco*RI/*Hind*III and *Eco*RI/*Pst*I double digests contain most of the ovalbumin sequences located to the left of the *Pst*I site (see the map in Fig. 2a). But, when the lengths of these 'a' fragments were compared, a striking anomaly became apparent. The *Eco*RI/*Hind*III fragment was shorter than the *Eco*RI/*Pst*I fragment by about 150 base pairs (2,600 and 2,750 base pairs, respectively). This shorter length cannot be attributed to a contaminating exonuclease activity in the *Hind*III enzyme, since all the restriction endonucleases used in this study were shown to be free of exonuclease activity even after prolonged incubations. Since the DNA fragment of an *Eco*RI digest which hybridises strongly to *Pst*A or *Hinf*A probes is 9,500 base pairs (Fig. 3, fragment 'a', lane 3), it follows that the *Eco*RI/*Hind*III and *Eco*RI/*Pst*I fragment 'a' must share the same *Eco*RI terminus. There is therefore a *Hind*III site in the chromosomal *Eco*RI/*Pst*I fragment 'a' at 150 base pairs from the *Pst*I terminus. But, no *Hind*III site has been found in the cloned ovalbumin ds cDNA or in the *in vitro* synthesised ds cDNA (see above and Fig. 2). One has therefore to conclude that the *Pst*I site present at one terminus of the DNA fragment 'a' of an *Eco*RI/*Pst*I double digest is not identical to the *Pst*I site present in the cloned or *in vitro* synthesised ds cDNA (Fig. 2). The simplest explanation to account for these results is that the DNA sequences complementary to the probes *Pst*A (or *Hinf*A) and *Pst*B (Fig. 2) are physically separated, and that the DNA sequence coding for the

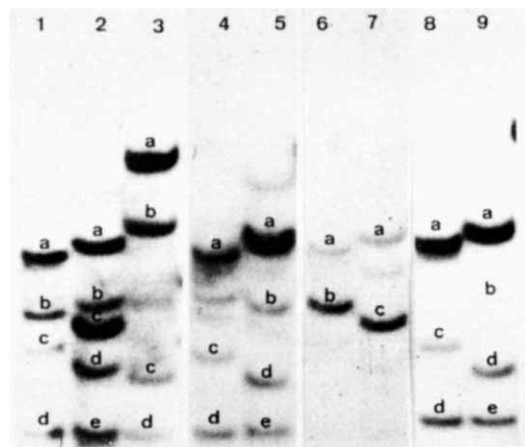


Fig. 4 Detection of fragments containing sequences complementary to defined parts of ovalbumin mRNA in chicken DNA cleaved with a combination of restriction enzymes. Electrophoresis was on a 1% Agarose gel. ^{32}P -labelled fragments of *Hha*ov were prepared by the method described in Fig. 3. To obtain probes containing only sequences complementary to ovalbumin mRNA, labelled *Hha*ov, *Pst*A and *Pst*B were annealed to a 10-fold excess of ovalbumin mRNA at 50.5°C in 100 mM PIPES buffer, pH 7.9, containing 70% formamide, 100 mM NaCl, 10 mM EDTA (after denaturation at 80°C in this buffer). DNA-DNA hybrids are not stable in these conditions¹⁷. Single-stranded plasmid sequences and GC tails were digested with S1 nuclease (Miles). The RNA-DNA hybrid was precipitated with ethanol and then treated with RNase A in 10 mM Tris-HCl pH 8.0, 1 mM EDTA at 50°C to generate single-stranded probes. All other methods were as described in Fig. 3. Filters were hybridised to single-stranded *Hha*ov (lanes 1-3), single-stranded *Pst*A (lanes 4 and 5), single-stranded *Pst*B (lanes 6 and 7) or to *Hinf*A (lanes 8 and 9). Lanes 1, 4, 6, 8, *Eco*RI/*Hind*III double digest of chicken DNA—sizes of the fragments are, in kilo base pairs: a, 2.6; b, 2.05; c, 1.8; d, 1.3. Lanes 2, 5, 7, 9, *Eco*RI/*Pst*I double digest of chicken DNA—sizes of the fragments are in kilo base pairs: a, 2.75; b, 2.15; c, 1.9; d, 1.65; e, 1.3. Lane 3, *Hind*III/*Pst*I double digest of chicken DNA—sizes in kilo base pairs: a, 4.4; b, 3.1; c, 1.5; d, 1.3. Fragment b does not hybridise to *Pst*A or *Hinf*A, but represents the main band with the *Pst*B probe (not shown).

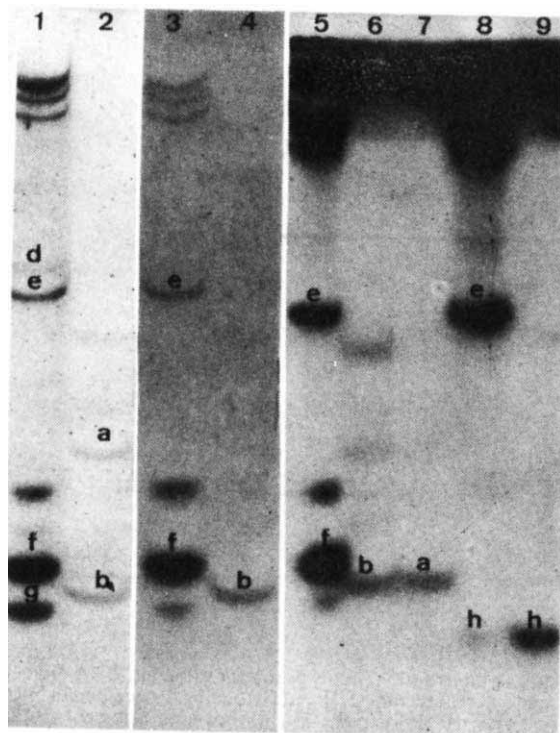


Fig. 5 Detection of ovalbumin gene sequences in chicken DNA cleaved with *HinfI* and *HinfI* plus *HaeIII* restriction nucleases. Chicken DNA was digested with endonuclease *HinfI* (gift of Dr D. Duck) or *HaeIII* (Biolabs), with SV40 DNA as internal marker, and electrophoresed on a 1.5% Agarose gel. All other procedures were as described in Fig. 3. Autoradiography was for 3 d. Filters were hybridised to either ^{32}P -*Hhaov* fragment, 10^8 c.p.m. μg^{-1} (lanes 1 and 2) or to ^{32}P -*HinfA* fragment 6×10^7 c.p.m. μg^{-1} (lanes 3–9). Lanes 1, 3 and 5, internal markers. Bands d and e are described in the legend to Fig. 1. Bands f and g correspond to the *HinfA* and *HinfB* fragments of *Hhaov*, 880 and 750 base pairs long, respectively (Fig. 2). Lanes 2, 4 and 6, chicken DNA digested with *HinfI*—sizes of the fragments in base pairs: a, 1,300; b, 800. Lane 7, chicken DNA digested with *HinfI* and *HaeIII*—size of band a, 800 base pairs. Lanes 8 and 9, internal markers; band e, see Fig. 1, fragment h, obtained by double digestion of *Hhaov* with *HaeIII* and *HinfI*, corresponds to the 650 base pair fragment between the *Hinf2* and *HaeIII* sites of the ds cDNA (see Fig. 2).

3'moiety of the mRNA (complementary to *PstA* and *HinfA*) is linked to a DNA sequence not coding for ovalbumin mRNA and containing a *HindIII* and a *PstI* site 150 base pairs apart.

Further evidence supporting the existence of a split ovalbumin gene was provided by the results of digestion of chicken DNA with *HinfI* endonuclease. As shown in Fig. 2, *HinfI* cuts the double-stranded DNA sequence complementary to ovalbumin mRNA in four places. One centrally located fragment, *HinfA*, is 880 base pairs long. This fragment should be found in a *HinfI* digest of chicken oviduct DNA, if the ovalbumin gene were not split in that region. Figure 5 shows the hybridisation pattern of the *HinfI* oviduct DNA fragments to ^{32}P -labelled *Hhaov* and *HinfA* probes. The *Hhaov* probe hybridised mainly to two DNA fragments, 800 and 1,300 base pairs (Fig. 5, lane 2), while *HinfA* hybridised only to the 800-base pair DNA fragment (Fig. 5, lanes 4 and 6). The use of a *HinfI* digest of the *Hhaov* fragment (Fig. 5, lanes 1, 3 and 5) as an internal marker for hybridisation and molecular weight determination demonstrates unambiguously that the 80-base pair difference between the length of the *HinfA* fragment and that of its chromosomal DNA counterpart is real. In addition, the 800-base pair chicken DNA *HinfI* fragment which hybridises to the *HinfA* probe was not cut by *HaeIII* restriction endonuclease as shown in Fig. 5, lane 7, where a *HaeIII*/*HinfI* double digest of chicken DNA was electrophoresed. The implication of these results is that there is an interruption in the chromosomal DNA sequence coding for ovalbumin mRNA between the *Hinf2* and *HaeIII* sites of the ds cDNA (see Fig. 2).

The ovalbumin gene is split in the coding region irrespective of the differentiated state of the cell

The data presented above show that in the chicken genome the DNA sequences coding for ovalbumin mRNA are split between the *Hinf2* and the *HaeIII* sites present in the ds cDNA. If the genomic DNA sequences coding for the remaining 5' part of the ovalbumin mRNA were in one piece one would expect that the *PstA* probe would hybridise to the same *EcoRI* or *HindIII* chicken DNA fragments which hybridise to the *PstB* probe, as these fragments would carry at least 320 bp also present in the *PstA* probe. As shown in Fig. 3, this was clearly not the case: the *EcoRI* and *HindIII* fragments 'b' which hybridised to the *PstB* probe did not hybridise to the *PstA* probe (Fig. 3). In consequence the genomic DNA sequences coding for ovalbumin mRNA must be split a second time near the *PstI* site of the ovalbumin ds cDNA (Fig. 2). The approximate location of the interruptions of the DNA sequences coding for ovalbumin mRNA are shown schematically in Fig. 6a, with respect to the restriction endonuclease sites present in the cloned ovalbumin ds cDNA (Fig. 2). Preliminary results suggest that the site of the first interruption lies close to the *HaeIII* restriction site. In any case at least one of the interruptions is located in a region coding for the ovalbumin protein: 1,161 nucleotides are required to code for the 387 amino acids of hen ovalbumin¹⁸, and it was recently shown that the first 250 nucleotides from the 3' end of the mRNA do not code for the protein¹⁵.

The occurrence of two interruptions could account for the origin of three of the fragments hybridising to the ovalbumin probe in an *EcoRI* digest of chicken DNA. The existence of a fourth band (Fig. 1) might suggest the presence of a third interruption.

Figure 6b and c show the map of the restriction endonuclease sites flanking regions I and III, which code for RNA sequences present in the 3' and in the 5' regions of the ovalbumin messenger, respectively. These maps were established from the lengths of the main fragments hybridising to the *PstA* and *PstB* probes (Figs 1, 3 and 4). We do not yet know the orientation of transcription of the ovalbumin RNA sequences. Comparison of these maps indicates that the DNA sequences surrounding region III (Fig. 6c) cannot be included in the 12,000-base pair DNA fragment containing region I (Fig. 6). Depending on the orientation of transcription, region I is separated in the genome from region III by either at least 2,500 base pair or at least 7,500 base pair (see Fig. 6b).

Our hybridisation studies do not reveal any DNA fragment which could be derived from an uninterrupted 'intact' ovalbumin gene. In addition we have no evidence for major variations in the structural organisation of the split gene. Since 80% of the laying hen oviduct cells are actively engaged in ovalbumin mRNA synthesis^{3,4} and since there is evidence that ovalbumin genes are present at approximately one copy per haploid genome (see ref. 3), we conclude that ovalbumin mRNA is transcribed from split genes. The oviduct DNA used in this study was isolated from total oviduct and it is therefore unlikely that loss of possible extrachromosomal elements carrying an 'intact' ovalbumin gene could have occurred. This peculiar arrangement of the genomic DNA coding for ovalbumin mRNA is apparently unrelated to the highly differentiated state of laying hen oviduct tubular cells. In all cases studied, the hybridisation patterns were identical for restriction digests of oviduct, erythrocyte or 5-d old embryo DNAs, indicating that there is no major modification of the ovalbumin gene organisation and environment during the differentiation of cells highly specialised in ovalbumin synthesis.

Conclusions

It is not known how the different DNA regions of the split ovalbumin gene are organised in chicken DNA. There are two types of model that would be consistent with our results. In the first type of model, the insertion model, the different regions I, II and III (Fig. 6a) are all on the same DNA segment in the same order and orientation as in the *in vitro* synthesised ds cDNA, but they are separated by DNA sequences (insertions) that are not represented in mature mRNA. In the second type of model, the

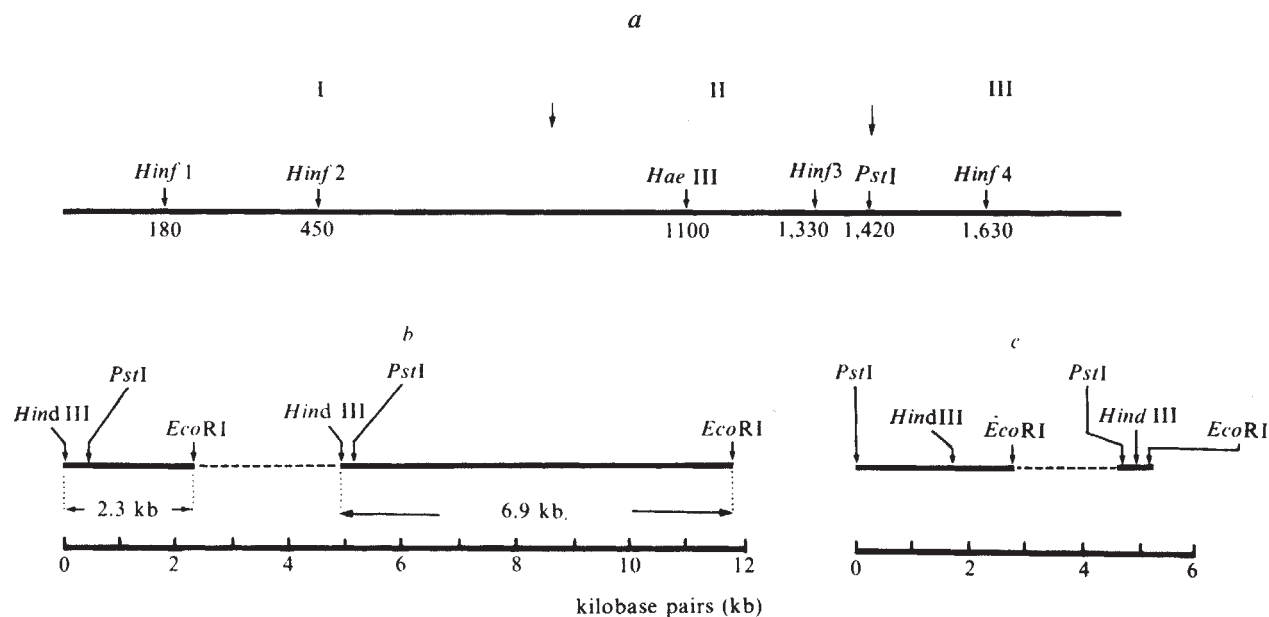


Fig. 6 Organisation of the sequences coding for ovalbumin mRNA in cellular DNA. *a*, Interruptions in the genomic ovalbumin DNA sequences. The arrows indicate the approximate location of the two interruptions and define three different regions (I, II and III) coding for ovalbumin mRNA. *b* and *c*, Physical map of restriction endonuclease cleavage sites flanking regions I and III, respectively. The dashed lines represent the DNA segments containing regions I or III. The sizes of the fragments used to establish these maps are (in kilo base pairs, see also Figs 1, 3 and 4) for region I: *EcoRI*, 9.5; *HindIII*, 4.9; *PstI*, 4.7; *EcoRI/HindIII*, 2.6; *EcoRI/PstI*, 2.75; *HindIII/PstI*, 4.4; for region III: *EcoRI*, 2.35; *HindIII*, 3.2; *PstI*, 4.7; *EcoRI/HindIII*, 2.05; *EcoRI/PstI*, 1.9; *HindIII/PstI*, 3.1. In each case the map was constructed using the first five values. From these maps one can predict sizes for the *HindIII/PstI* fragments which are in very good agreement with those found experimentally.

different regions are not in the same order and orientation as in the ds cDNA and could even be on different chromosomes. Further mapping of the ovalbumin sequences in chicken DNA is in progress, in the hope of finding a fragment which will contain all three regions of the ovalbumin gene, thereby supporting the insertion model. Up to now, no such fragment has been found. This does not rule out the insertion model, since the chance of finding such a fragment decreases with the length of the insertion.

The existence of DNA insertions within eukaryotic genes has recently been demonstrated for 28S ribosomal cistrons of *Drosophila melanogaster*¹⁹⁻²². But, since intact *Drosophila* ribosomal cistrons were also found, the physiological relevance of the insertion could be questioned. A 93-base pair insertion was detected and sequenced in the 5' region of the protein coding sequence of a cloned immunoglobulin variable gene of mouse embryo (S. Tonegawa, personal communication), and a 1,250-base pair insertion was found within the coding sequence of a cloned mouse plasmacytoma λ chain DNA, separating the sequences coding for the variable and constant regions of the λ chain (C. Brack & S. Tonegawa, personal communication). Using an approach very similar to ours, which eliminates the possible complications related to accidental selection of defective genes and/or DNA rearrangements during cloning, A. Jeffreys and R. Flavell (*Cell*, in the press) have detected the presence of a 600-base pair insertion within the protein sequence of the rabbit β -globin gene. A similar insertion was found in a cloned mouse β -globin gene (P. Leder, personal communication). It seems therefore that the splitting of genes within the protein coding sequence may have some generality in eukaryotic cells.

It is too early to speculate on the role of the split gene organisation in the regulation of gene expression at the transcriptional and/or post-transcriptional levels. But, the two types of model discussed above for the ovalbumin gene lead to different predictions concerning the general processes by which a split gene could be transcribed to give a single continuous mRNA. In the case of the insertion model three possible mechanisms already proposed to account for the mosaic late adenovirus mRNAs²³⁻²⁵ could apply to the ovalbumin gene: (1) RNA polymerase B which is responsible for ovalbumin mRNA synthesis⁴ could jump with the attached nascent RNA across looped out insertions, transcribing in an ordered fashion the ovalbumin gene regions III, II

and I (Fig. 6a) to yield directly the mRNA; (2) both insertions and messenger coding regions could be transcribed resulting in an ovalbumin mRNA precursor. The mature mRNA could then be formed by excision of the insertion transcripts and splicing of the mRNA sequences. The size of the precursor should be at least twice that of mature ovalbumin mRNA, since the minimum distance separating regions I and III in the insertion model is about 2,500 base pairs (see above and Fig. 6); (3) each of the mRNA coding segments may be transcribed independently and the ovalbumin mRNA generated by ligation of the transcripts. This latter mechanism, but not the two former ones, could also apply to the second type of model of the split gene where the mRNA coding regions are not placed in an ordered fashion (see above). We have no direct evidence to distinguish between these three possibilities, since the previous failure to detect an ovalbumin mRNA precursor²⁶ must be reinvestigated in the light of the recent finding concerning the short half-life of the β -globin mRNA precursor^{27,28}.

The split gene type of organisation poses additional problems for studying *in vitro* the mechanisms involved in the specificity of transcription. The phenomenon of split genes in eukaryotic cells also diminishes the chance of obtaining single clones containing the DNA sequences coding for the entire mRNA and its putative regulatory elements. Furthermore, even where such clones may be isolated, it is unlikely that they could be used for the expression of eukaryotic genes in bacteria, as the necessary machinery for post-transcriptional processing may well not be present in prokaryotic organisms. Progress in this field may have to be based on the linking of known bacterial regulatory elements to ds cDNA clones containing the entire protein coding sequence.

We thank J. Sümegi, J. P. LePennec and A. Krust for generous gifts of materials, and M. C. Chanal for technical assistance. This work was supported by grants from INSERM (CRT 76.5.468 and 76.5.462), CNRS (ATP 2117) and the Fondation pour la Recherche Médicale Française. R. Breathnach was supported by a long-term EMBO fellowship.

Received 11 October; accepted 17 October 1977.

1. Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3628-3632 (1976).
2. Oka, T. & Schimke, R. T. *J. cell. Biol.* **43**, 123-132 (1969).
3. Palmiter, R. D. *Cell* **4**, 189-197 (1975).

4. Bellard, M., Gannon, F. & Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42**, (in the press).
5. Botchan, M., Topp, W. & Sambrook, J. *Cell* **9**, 269–287 (1976).
6. Southern, E. M. *J. molec. Biol.* **98**, 503–518 (1975).
7. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
8. Humphries, P. *et al. Nucleic Acids Res.* **4**, 2389–2406 (1977).
9. Monahan, J. J., Woo, S. L. C., Liarakos, C. D. & O'Malley, B. W. *J. biol. Chem.* **252**, 4722–4728 (1977).
10. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
11. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 32–38 (1973).
12. Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641–646 (1966).
13. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1189 (1975).
14. Philipson, L., Peterson, U. & Lindberg, U. *Molecular Biology of Adenovirus. Virology Monographs* **14**, (Springer, Berlin, 1975).
15. Brownlee, G. G. & Cartwright, E. M. *J. molec. Biol.* **114**, 93–117 (1977).
16. Maxam, A. A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
17. Casey, J. & Davidson, N. *Nucleic Acids Res.* **4**, 1539–1552 (1977).
18. Woo, S. L. *et al. J. biol. Chem.* **250**, 7027–7039 (1975).
19. Glover, D. M. & Hogness, D. S. *Cell* **10**, 167–176 (1977).
20. White, R. L. & Hogness, D. S. *Cell* **10**, 177–192 (1977).
21. Wellauer, P. K. & Dawid, I. B. *Cell* **10**, 193–212 (1977).
22. Pellegrini, M., Manning, J. & Davidson, N. *Cell* **10**, 213–224 (1977).
23. Berget, S. M., Moore, C. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171–3175 (1975).
24. Sambrook, J. *Nature* **268**, 101–104 (1977).
25. Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42**, (in the press).
26. McKnight, G. S. & Schimke, R. T. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4327–4331 (1974).
27. Bastos, R. N. & Aviv, H. *Cell* **11**, 641–650 (1977).
28. Curtis, P. J., Mantei, N., Van den Berg, J. & Weissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3184–3188 (1977).

letters to nature

Low energy γ -ray observation of NGC4151

OBSERVATIONS of the Seyfert galaxy NGC4151 by Ariel V, OSO VII and UCSD (refs 1–3 respectively) have demonstrated that it has a relatively flat X-ray spectrum up to photon energies greater than 100 keV. The X-ray luminosity of this object is of the order of 4×10^{43} erg s $^{-1}$ if a distance of 20 Mpc is assumed, and exceeds the integrated luminosity at all the other observed greater wavelengths. Long-term observations⁴ by Ariel V indicate that the X-ray emission may originate from a compact object with dimensions less than 8×10^{15} cm, presumably the nucleus of the Galaxy. Here we present the preliminary results of a positive measurement in the low energy γ -ray region of the spectrum, which are in good agreement with the X-ray data

and confirm the intense high energy luminosity of the object.

The Seyfert galaxy NGC4151 was one of the celestial objects observed during a recent balloon flight, from Palestine Texas on 22 and 23 May 1977, of the MISO low energy (0.2–20 MeV) γ -ray telescope^{5,6}. The instrument is a semiactively-shielded Compton coincidence system having a sensitive area of about 500 cm 2 and an aperture of $3^\circ \times 20^\circ$ FWHM. The gondola was steered by an orientation system designed to point the telescope in both zenithal and azimuthal planes. A hard X-ray telescope operating in the 30–250 keV energy range was also incorporated in the scientific payload. This detector viewed the same region of the sky as the γ -ray device, with an aperture of $2^\circ \times 5^\circ$ FWHM and was designed around a sodium iodide crystal 3 mm thick and having a sensitive area of about 70 cm 2 .

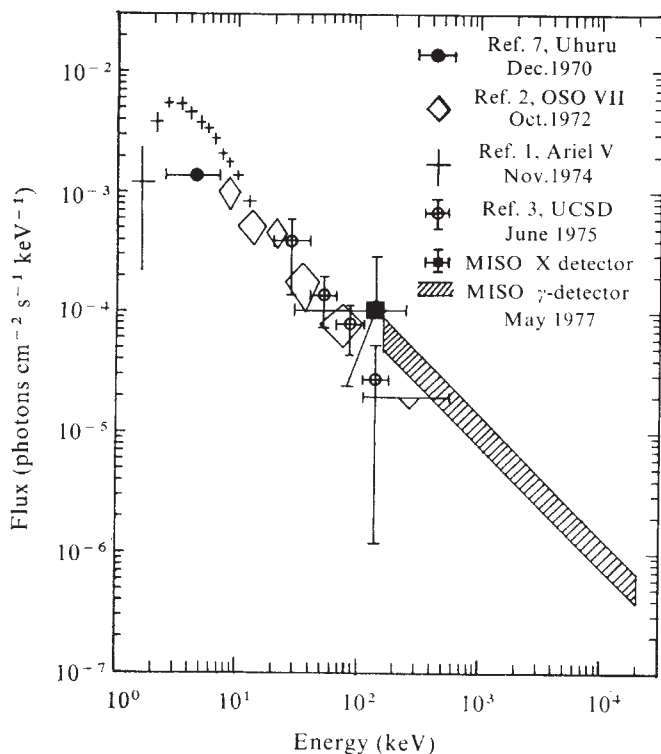
The entire observation of NGC4151 was made within 18° of the zenith at a floating altitude of about 7 mbar over a period of 3 h. Half of this time was used to obtain the background evaluation. NGC4151 was detected at the 6.5σ level (based on the errors of the differences of the on and off source counting rate) in the 0.15–20 MeV energy band. In this preliminary analysis only real time counting rate data, processed by an on-line computer, from the sodium iodide crystal (component of the Compton coincidence system) were used. No spectral information is available at this stage. The observed counting rates after background subtraction at the same zenith angle were corrected for any changes in float altitude. No significant variation in background counting rate with azimuthal angle was observed.

Four transits of the source through the fine aperture of the telescope collimator were performed and a positive increase in γ -ray events observed in each case. The excess photon counting rate of 8.4 ± 1.3 counts s $^{-1}$ at γ -ray energies and 0.31 ± 0.23 counts s $^{-1}$ at X-ray wave lengths has been interpreted as a continuation of the previously measured hard X-ray spectrum. The γ -ray excess has been converted to a photon spectrum at the top of the atmosphere using conversion factors calculated for an input spectral index -1 .

The γ flux obtained in this way is shown in Fig. 1 together with relevant soft and hard X-ray data. The hatched area takes into account both the uncertainties derived from statistical fluctuations at 1σ level and the lack of knowledge of the precise exposure of the source. This latter uncertainty is caused by not having continuously updated orientation data in the quick-look mode to better than 1° .

The γ -ray results show good agreement with a power law extension of the X-ray data using a spectral index similar

Fig. 1 The soft and hard X-ray measurements of NGC4151 compared with the MISO X and γ measurements. The integral γ -ray counting rate has been converted into a differential spectrum (hatched area) using a spectral index -1 .



to that suggested by Baity *et al.*². The extension of the emission spectrum of NGC 4151 up to 20 MeV would correspond to a luminosity of $(8 \pm 2) \times 10^{45}$ erg s⁻¹ for a distance of 20 Mpc.

Laboratorio TESRE, CNR,
Bologna, Italy

Laboratorio di Fisica Cosmica e Tecnologie Relative,
CNR, Milano, Italy

Department of Physics,
University of Southampton,
Southampton, UK

G. DI COCCO

G. BOELLA
F. PEROTTI
R. STIGLITZ
G. VILLA

R. E. BAKER
R. C. BUTLER
A. J. DEAN
S. J. MARTIN
D. RAMSDEN

Received 29 July; accepted 3 October 1977.

1. Ives, J. C., Sanford, P. W. & Penston, M. V. *Astrophys. J. Lett.* **207**, L159–L162 (1976).
2. Baity, W. A., Jones, T. W., Wheaton, Wm. A. & Peterson, L. E. *Astrophys. J. Lett.* **199**, L5–L8 (1975).
3. Paciesas, W. S., Mushotzky, R. F. & Pelling, R. M. *Mon. Not. R. astr. Soc.* **178**, 23P–25P (1977).
4. Elvis, M. *Mon. Not. R. astr. Soc.* **177**, 7P–12P (1976).
5. Maccagni, D., Paizis, C., Perotti, F., Stiglitz, R., Villa, G., Bailey, P., Dean, A. J. & Ramsden, D. *Proc. 9th Esab Symp. Frascati* 185–193 (1974).
6. Maccagni, D., *et al.* *14th Int. Cosmic Ray Conf. Munich* **9**, 3135–3140 (1975).
7. Gursky, H., Kellog, E. M., Leong, C., Tananbaum, H. & Giacconi, R. *Astrophys. J. Lett.* **165**, L43–L48 (1971).

Optical candidates for two X-ray bursters and an X-ray pulsar

WE suggest faint ($B \sim 18$), blue stars as the optical counterparts of two X-ray bursters, 4U1636–53 (MXB1636–53)^{1–4} and 4U1735–44 (MXB1735–44; KGX345–6?)^{1,4–6}, and for the 7-s X-ray pulsar, 4U1626–67^{1,7,8}. The candidate stars have large ultraviolet excesses and were discovered well within the 20" and 30" X-ray error radii determined using the rotating modulation collimator experiment aboard the SAS-3 X-ray Observatory^{9,10}. Photographic observations of several X-ray error boxes were performed (by C.R.C. and W.A.H.) in April 1976 (before final SAS-3 X-ray positions were known) with the 4-m telescope at the Cerro Tololo Interamerican Observatory (CTIO). R , B and U plates were obtained for the regions including 4U1636–53 and 4U1735–44. The R plates (RG610 filter with 127-04 emulsion) are shown as

finding charts in ref. 9. Preliminary iris photometry of some of the plates was performed to identify the bluest stars in each error box for further study.

The photoelectric photometry was performed (by J.E.M.) using the 1.5-m telescope at CTIO during 8–13 June, 1977. A 10" aperture, a standard UBV filter set and an RCA 4516 photomultiplier tube were used during the entire run. For each colour typical integration times were 10 s for standard stars and 50 s for the programme stars and sky background. All objects were observed in the colour sequence $BVUB$ as a check on sky conditions, telescope tracking and the stability of the instrumentation. The observations were made between the fourth quarter and new moon. The sky conditions were of photometric quality throughout each night with the exception of 9 June, which had some cirrus. Five or six UBV standard stars¹¹ were observed several times each night. A television guider, mounted at the customary position of the offset-guider eyepiece, was used on four nights (8, 9, 12 and 13 June), and provided the crucial capability to centre stars as faint as $B \sim 20$ in the photometer aperture.

About 100 faint ($B \sim 15$ –20) stars in the fields of 16 X-ray sources were examined during the six nights. Most of the observations were made within the 20–30" radius error circles determined using SAS-3 data. Stars with large ultraviolet excesses were discovered in the fields of three X-ray sources: 4U1626–67, 4U1636–53 and 4U1735–44.

A summary of the photometry data, coordinates of the stars and references to literature containing finding charts are given in Table 1. A measurement of the candidate for 4U1626–67 made while setting up on 12 June (0553 UT) and not included in Table 1 indicates that the star may be variable on timescales of minutes, but further observations are necessary to establish this. The three measurements of the candidate for 4U1735–44 show no evidence of variability during the course of four nights. Iris photometry of two R plates taken four days apart in April 1976 also shows no variability. A spectrogram of this candidate obtained by Bond¹² shows a generally featureless blue continuum with weak $H\beta$ emission. HeII λ 4,686 and HeI λ 5,875 may also be present weakly in emission.

Figure 1 is a colour-colour diagram showing the measured locations of the candidate stars and several established optical counterparts (see refs 13–17), and the calculated location of a model accretion disk around a massive black hole¹⁸. The candidates for 4U1626–67 and 4U1636–53 lie further from the reddened supergiant or main sequence than any known counterparts and have colours inconsistent with those of a reddened, normal star. The candidate for 4U1735–44 has the colours of Sco X-1 as well as the colours of a late O or B0 star suffering ~ 1 mag of extinction.

Table 1 Description of optical candidates

X-ray source	Star no.*	Star position ($\pm 3''$)		UT (June 1977)	$V^\dagger$	$B-V^\dagger$	$U-B^\dagger$
		RA (1950) ($h^m, s^{.11}$)	Dec. (1950)				
4U1626–67 (2S1627–673)	4	16h27m14.2s (321.8–13.1)	–67°21'16"	12d06h05m 12 06 34	18.67 18.42	+0.00 +0.26	–1.18 –1.23
4U1636–53 (2S1636–536)	3	16 36 56.2 (332.9–4.8)	–53°39'15"	12 03 21	17.52	+0.69	–0.70
4U1735–44 (2S1735–444)	5	17 35 19.0 (346.1–7.0)	–44°25'19"	09 06 39 11 05 10 13 06 54	17.43 17.40 17.52	+0.24 +0.22 +0.21	–0.78 –0.86 –0.81

*Finding charts showing the SAS-3 (2S) positions and labelled with these star numbers are given for 4U1626–67 (in ref. 10) and for 4U1636–53 and 4U1735–44 (in ref. 9).

† Uncertainties ($\sim 90\%$ confidence) in V , $B-V$ and $U-B$ are + 0.15 mag. for 4U1626–67 and + 0.10 mag. for 4U1636–53 and 4U1735–44. Uncertainties due to counting statistics are $\lesssim \pm 0.05$ mag.

The candidates for 4U1735-44 and 4U1636-53 bring the number of possible stellar counterparts to X-ray burst sources to three (the third is Davidson's candidate for Ser X-1^{9,20}). The colours of the two presented here are far from those estimated by Katz¹⁸ for an accretion disk surrounding a massive black hole. Evidence of the binary nature of any or all these candidates could resolve the uncertainty over the nature of the bursters (see refs 21, 22). The extreme colours of the candidate for 4U1626-67 are remarkable since this source is a rapid X-ray pulsar⁸. The optical emission of all other X-ray pulsars, with the exception of HZ Her, is dominated by that of the OB companion.

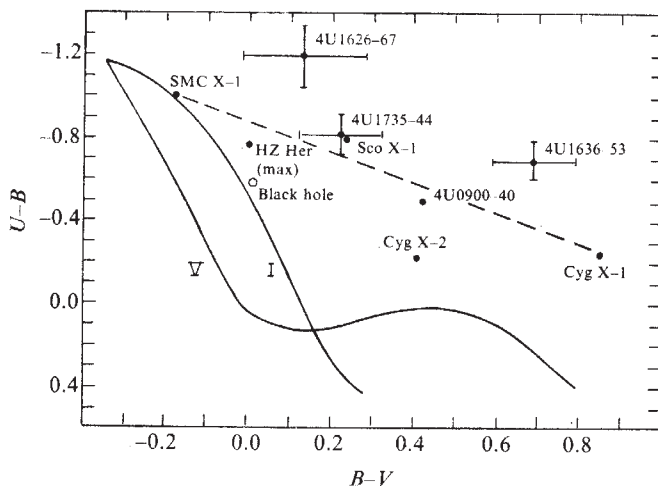


Fig. 1 A colour-colour diagram in the UBV system showing the three candidates presented in this paper, several well-established optical counterparts¹³⁻¹⁷ and a model accretion disk surrounding a massive black hole¹⁸. The solid curves give the locus of colours for unreddened, main-sequence stars (V) and supergiants (I) following Davis.²⁴ The dashed curve is the reddening line for a B0 supergiant.

Possibly 4U1626-67 is an even more extreme example of the HZ Her phenomenon, in which the optical luminosity at maximum light is a direct result of the X-ray emission (see ref. 23).

We thank G. Clark, E. Ng, R. Kelly and other members of the SAS-3 group, E. Cosgrave and the staff at CTIO, and M. Liller, R. McCrosky, C. Shao and the Harvard College Observatory for use of their iris photometer and astrometry machine. This work was supported in part by NASA (grant NGR 23-005-465, Contract NAS-5-11450) and by MIT. J.E.M., C.R.C. and W.A.H. are visiting astronomers, Cerro Tololo Inter-American Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under contract with the USNSF.

J. E. MCCLINTOCK
C. R. CANIZARES
H. V. BRADT
R. E. DOXSEY
J. G. JERNIGAN

Department of Physics and Center for Space Research,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

W. A. HILTNER

Department of Astronomy,
University of Michigan,
Ann Arbor, Michigan 48104

Received 30 August; accepted 6 October 1977.

1. Forman, W. *et al. Astrophys. J. Suppl.* (in the press).
2. Swank, J. H., Becker, R. H., Pravdo, S. H., Saba, J. R. & Serlemitsos, P. J. *IAU Circ. No. 3000* (1976).

3. Hoffman, J. A., Lewin, W. H. G. & Doty, J. *Astrophys. J. Lett.* (submitted).
4. McClintock, J. E. *IAU Circ. No. 3084* (1977).
5. Sagdeev, R. Z. & Melioranskij, A. S. *IAU Circ. No. 2959* (1976).
6. Lewin, W. H. G., Hoffman, J. A., Doty, J., Li, F. K. & McClintock, J. E. *IAU Circ. No. 3075* (1977).
7. McClintock, J. E. *IAU Circ. No. 3088* (1977).
8. Rappaport, S. *et al. Astrophys. J. Lett.* (submitted).
9. Jernigan, J. G., Apparao, K. M. V., Bradt, H. V., Doxsey, R. E. & McClintock, J. E. *Nature* **270**, 321-323 (1977).
10. Bradt, H. V. *et al. Nature* **269**, 496-497 (1977).
11. Landolt, A. U. *Astr. J.* **78**, 959-981 (1973).
12. Bond, H. E. *IAU Circ. No. 3085* (1977).
13. Hiltner, W. A. & Mook, D. E. *A. Rev. Astr. Astrophys.* **8**, 139-160 (1970).
14. Murdin, P. & Webster, B. L. *Nature* **233**, 110 (1971).
15. Hiltner, W. A., Werner, J. & Osmer, P. *Astrophys. J. Lett.* **175**, L19-L22 (1972).
16. Webster, B. L., Martin, W. L., Feast, M. W. & Andrews, P. J. *Nature phys. Sci.* **240**, 183 (1972).
17. Boynton, P. E., Canterna, R., Crosa, L., Deeter, J. & Gerend, D. *Astrophys. J.* **186**, 617-624 (1973).
18. Katz, J. I. *Astrophys. Lett.* **18**, 69-72 (1977).
19. Davidson, A. *IAU Circ. No. 2824* (1975).
20. Doxsey, R. E., Apparao, K. M. V., Bradt, H. V., Dower, R. G. & Jernigan, J. G. *Nature* **269**, 112-116 (1977).
21. Lewin, W. H. G. *et al. Nature* **267**, 28-30 (1977).
22. Ostriker, J. P. paper presented at 8th Texas Symp. Relativistic Astrophys. Boston (1976).
23. Joss, P. C., Avni, Y. & Rappaport, S. *Astrophys. J.* (submitted).
24. Davis, R. J. *Astrophys. J.* **213**, 105-110 (1977).

Positions of three X-ray burst sources

PRECISE (20-30'') positions of three steady X-ray sources which have been identified recently as X-ray burst sources 4U1636-53 = MXB1636-53 (refs 1, 2), 4U1728-33 = MXB1728-34 (refs 3, 4), and 4U1735-44 = MXB1735-44 (refs 5, 6, 7) are reported here. These positions, hereafter designated with a 2S prefix, have been derived from data obtained with the SAS-3 rotating modulation collimators during a survey of the galactic plane⁸⁻¹⁰. The small solid angles of the error regions make possible a thorough survey of all potential optical, infrared, and radio counterparts of these burst sources. Two of these positions, 2S1636-536 and 2S1735-444, have led to the probable identifications of optical counterparts¹¹.

The burst sources MXB1636-53 and MXB1728-34 were discovered respectively with the OSO-8 (ref. 1) and SAS-3 (refs 3, 12) satellites. The source MXB1735-44, located precisely⁵ with SAS-3, is probably the burst source KGX345-6 that was discovered earlier^{6,7} with the COSMOS 428 satellite. All three sources emit X-ray bursts with peak flux densities (2-15 keV) comparable to, or greater than, that of the Crab nebula and with rise times of 1-2 s.

The association of the X-ray burst sources with the previously known Uhuru sources is now well established (Fig. 1). These identifications have been made possible through the determination^{2,3,5} of the burst source positions to an accuracy of 6-10' with the SAS-3 slat collimators¹³. Positions of the steady components of these sources have been measured with the Uhuru¹⁴, Copernicus¹⁵, and Ariel V¹⁶ satellites. The present SAS-3 positions are shown in Figs 1 and 2 and Table 1. Astrometric positions and stellar magnitudes are given in Table 2.

An optical study based upon the Copernicus position for MXB1636-53 led to the suggested optical identification¹⁷ of a flare star (star M in Fig. 2) that lies outside but near the 2S error region. A recent study of the 2S1636-536 and 2S1735-444 positions led to the discovery¹¹ of two faint $V \sim 17.5$, ultraviolet excess stars that lie well within the 2S regions (stars 3 and 5 respectively in Fig. 2). A search¹⁸ for an optical counterpart for MXB1728-34 has yielded no likely candidates. Several bright ($K < 10$ mag) infrared sources have been detected, however, within $\sim 1'$ of 2S1728-337 by I. Glass and by S. Kleinmann (personal communications).

Altogether, we have measured with SAS-3 precise (20-35'') positions of six steady X-ray sources (other than known globular clusters) which have been identified^{1-7,13,19-22} as

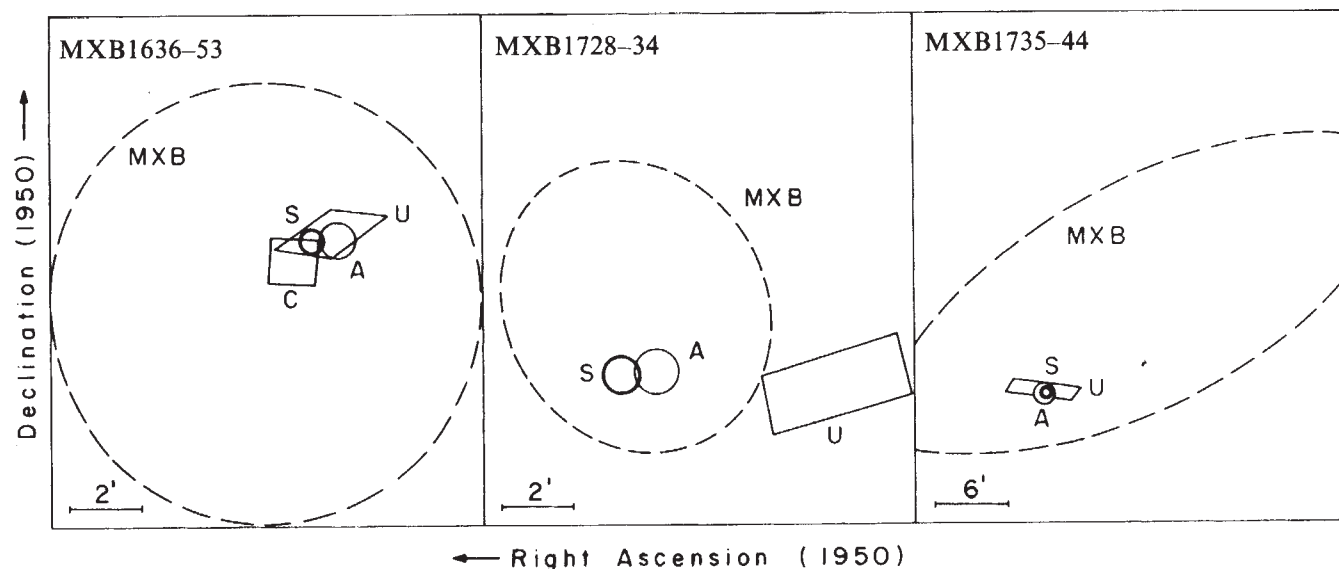


Fig. 1 The burst source locations (dashed lines) and steady source locations (solid lines) MXB1636-53; MXB1728-34 and MXB1735-44. The burst source locations designated MXB, were obtained with the SAS-3 slat collimators (refs 2, 3, 5 and J. Hoffman, personal communication). The steady source locations are from the Uhuru 4U catalogue (U, ref. 14), the Ariel V rotating modulation collimator (A, ref. 16), the Copernicus satellite (C, ref. 15), and the SAS-3 rotating modulation collimator [S (dark circles) present work].

counterparts of X-ray bursters. Previously we reported⁹ the positions of Ser X-1 (2S1837+049), A1905+00 (2S1905+000), and 4U1915-05 (2S1916-053). All six lie within 40° of the galactic centre and within 10° of the galactic plane. The intensities of the steady sources measured with SAS-3 agree roughly with those reported previously^{14,23}. None of these sources is known to exhibit periodicities, eclipses, or rapid variability other than bursts. These featureless characteristics are typical of the X-ray sources which are located in the galactic bulge. We also note that the burst sources exhibit a galactic distribution²⁴ similar to the bulge sources.

The bulge X-ray sources are similar to the ~ 6 steady X-ray sources which are identified^{23,25-27} with globular clusters, in

that they exhibit a deficiency of hard X rays and are associated with late-type stars in regions of high stellar density^{28,29}. Four globular clusters, NGC6624 (ref. 30), NGC1851 (refs 31, 32), NGC6441 (ref. 33) and Liller 1 (refs 34, 35) are now believed to be X-ray bursters. The physical significance of these apparent similarities among burst sources, the bulge sources, and globular cluster sources is not well understood (see ref. 29). This makes the search for and study of optical, infrared, and radio counterparts of the burst sources particularly compelling.

We acknowledge the support of the staffs of the Center for Space Research at MIT and the Goddard Space Flight Center of NASA. We thank C. Canizares, R. Dower, J. Kulik, W. Lewin, and C. Reid for helpful conversations and assistance.

Fig. 2 Finding charts for the SAS-3 positions. The charts for 2S1636-536 and 2S1735-444 are from R plates taken with the CTIO 4 m telescope by W. A. Hiltner and C. Canizares. The chart for 2S1728-337 is from the southern extension of the Palomar Sky Survey (National Geographic Society). North is up and east is to the left. The flare star designated M (2S1636-536) has been suggested¹⁷ as an optical counterpart. The 2S1636-536 and 2S1735-444 regions each contain a star (numbers 3 and 5 respectively) with a pronounced ultraviolet excess¹¹.

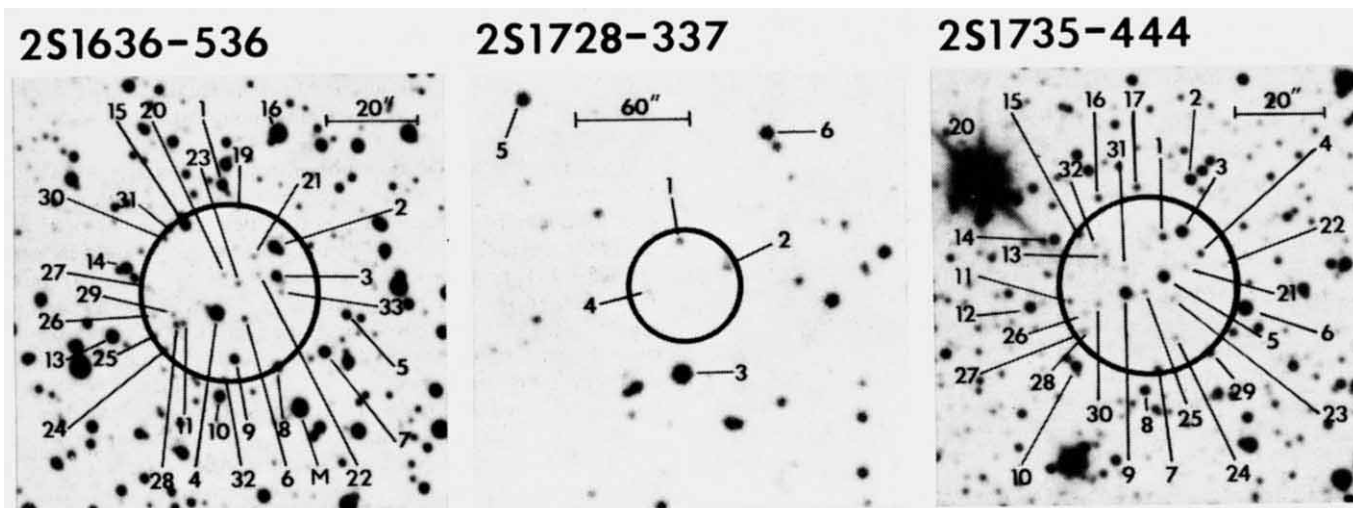


Table 1 Celestial positions

SAS-3 designation	Other designations*	Position (1950)		l^{II} b^{II}	Error radius (90%)	Flux density† (2–11 keV)	Comments
		α	δ				
2S1636–536	4U1636–53	16h 36min 57.6s 249.2400°	–53°39'21'' –53.6558°	332.9° –4.8°	20''	240 μ Jy	MXB1636–53
2S1728–337	GX354+0 4U1728–33	17 28 39.6 262.1650	–33 47 52 –33.7978	354.3 –0.2	30	95	MXB1728–34
2S1735–444	GX346–7 4U1735–44	17 35 19.5 263.8313	–44 25 22 –44.4228	346.1 –7.0	20	160	MXB1735–44 KGX345–6?

*Refs 14, 36, 37.

†1.0 μ Jy corresponds to 2.2×10^{-11} erg cm $^{-2}$ s $^{-1}$ (2–11 keV). $I_{\text{Crab}} = 1,060$ μ Jy. See refs 8, 9.

Table 2 Stellar astrometry and magnitudes

Source	Star	m_v^*	Position (1950)†	
			α	δ
2S1636–536	4	16.2		
	9	17.8		
	16		16 h 36 min 56.3 s	–53° 38' 45''
2S1728–337	2	18.2		
	3	13.5	17 28 39.7	–33 48 36
	5		17 28 46.1	–33 46 11
	6		17 28 35.9	–33 46 30
2S1735–444	3	17.6		
	20		17 35 22.9	–44 24 59

* From photoelectric photometry $^{11} \pm 0.1$ mag (± 0.25 mag for star 2).† Precise to $\leq 3''$.

27. Markert, T. H., Backman, D. E., Canizares, C. R., Clark, G. W. & Levine, A. M. *Nature* 257, 32–33 (1975).
 28. Canizares, C. *Astrophys. J.* 201, 589–592 (1975).
 29. Markert, T. H. *et al. Astrophys. J.* (in the press).
 30. Grindlay, J. *et al. Astrophys. J. Lett.* 205, L127–L130 (1976).
 31. Forman, W. & Jones, C. *Astrophys. J. Lett.* 207, L177–L180 (1976).
 32. Clark, G. W. & Li, F. K. *IAU Circ. No.* 3092 (1977).
 33. Li, F. K. & Clark, G. W. *IAU Circ. No.* 3095 (1977).
 34. Liller, W. *IAU Circ. No.* 2929 (1976).
 35. Lewin, W. H. G. *et al. Astrophys. J. Lett.* 207, L95–L99 (1976).
 36. Kellogg, E., Gursky, H., Murray, S., Tananbaum, H. & Giacconi, R. *Astrophys. J. Lett.* 169, L99–L103 (1971).
 37. Heinz, C. *et al. IAU Circ. No.* 2466 (1972).

Identification of the $\lambda 2,200\text{\AA}$ interstellar absorption feature

A BROAD absorption feature centred on $\lambda 2,200\text{\AA}$ with a half-width of $\sim 300\text{\AA}$ appears in the spectra of reddened stars $^{1-3}$. This conspicuous feature in the interstellar extinction curve, might hold an important clue to the identity of a major component of interstellar matter, but it has defied identification for over a decade. Here we identify this band as representing the integrated effect of a set of bicyclic compounds, each with the empirical formula $\text{C}_8\text{H}_6\text{N}_2$. Such nitrogenated structures could form in stellar mass flows of the type which we have also discussed 4 . A significant mass fraction of all interstellar material might exist in this form.

Graphite particles have been considered the most plausible candidate for the $\lambda 2,200\text{\AA}$ absorption feature. Whilst a small particle resonance in graphite can occur close to $\lambda 2,200\text{\AA}$, the central wavelength of this feature is sensitively dependent on particle shape 5 . Spherical particles with sizes small compared to the wavelength are necessary to produce agreement with observational data, but a more realistic distribution of shapes would produce a considerably broader absorption feature than is required. It therefore seems that a narrower molecular absorption must be superimposed on an underlying broader extinction hump which could be caused by extinction from graphite grains with a wide spread in their shapes.

We have discussed a possible molecular origin for the $\lambda 2,200\text{\AA}$ band due to $\pi \rightarrow \pi^*$ electronic transitions in a wide class of molecules involving conjugated double bonds 6 . We now limit our search to the nitrogenated heterocyclic compounds listed in Table 1. The first four compounds represent all possible arrangements of two N atoms in a bicyclic structure, with the hetero-atoms confined to one ring only. The fifth compound is an isomer where there is one N atom in each of the two rings.

An average molar absorptivity function $\langle \epsilon(\lambda) \rangle$ was computed for these materials from available spectroscopic data 7 . A normalised absorptivity $A(\lambda)$ given by

$$A(\lambda) = \frac{\langle \epsilon(\lambda) \rangle - \langle \epsilon(\lambda_0) \rangle}{\langle \epsilon(\lambda_1) \rangle - \langle \epsilon(\lambda_0) \rangle} \quad (1)$$

with $\lambda_0^{-1} = 3.8 \mu\text{m}^{-1}$, $\lambda_1^{-1} = 4.55 \mu\text{m}^{-1}$ is plotted in Fig. 1. Our computed curve for $\text{C}_8\text{H}_6\text{N}_2$ isomers agrees exactly with the interstellar extinction data with respect to the central wavelength, but the 'average' interstellar band is apparently $\sim 30\%$ wider. The latter departure could easily be ascribed to an

This work was supported in part by NASA. This article is the fourth in a series on positions of galactic X-ray sources obtained with SAS-3.

J. G. JERNIGAN
 K. M. V. APPARAO*
 H. V. BRADT
 R. E. DOXSEY
 J. E. MCCLINTOCK

Department of Physics and Center for Space Research,
 Massachusetts Institute of Technology,
 Cambridge, Massachusetts 02139

Received 23 August; accepted 22 September 1977.

* Permanent address: Tata Institute of Fundamental Research, Bombay, India.

- Swank, J. H., Becker, R. H., Pravdo, S. H., Saba, J. R. & Serlemitsos, P. J. *IAU Circ. No.* 3000 (1976).
- Hoffman, J. A., Lewin, W. H. G. & Doty, J. *Astrophys. J. Lett.* (in the press); *IAU Circ. No.* 3025 (1977).
- Hoffman, J. A. *et al. Astrophys. J. Lett.* 210, L13–L17 (1976).
- Carpenter, G. F., Skinner, G. K., Wilson, A. M. & Willmore, A. P. *Nature* 262, 473–474 (1976).
- Lewin, W. H. G., Hoffman, J. A., Doty, J., Li, F. K. & McClintock, J. E. *IAU Circ. No.* 3075 (1977).
- Babushkina, O. P. *et al. Sov. Astr. Lett.* 1, 32–34 (1975).
- Sagdeev, R. Z. *IAU Circ. No.* 2959 (1976).
- Bradt, H. V. *et al. Nature* 269, 21–25 (1977).
- Doxsey, R. E., Apparao, K. M. V., Bradt, H. V., Dower, R. G. & Jernigan, J. G. *Nature* 269, 112–116 (1977).
- Bradt, H. V. *et al. Nature* 269, 496–497 (1977).
- McClintock, J. E., Canizares, C., Bradt, H. V., Doxsey, R., Jernigan, J. G. & Hiltner, W. A. *H. Nature* 270, 320–321 (1977); *IAU Circ. Nos* 3084, 3088 (1977).
- Lewin, W. H. G. *IAU Circ. No.* 2922 (1976).
- Lewin, W. H. G. *et al. Mon. Not. R. astr. Soc.* 177, 93p–100p (1976).
- Forman, W. *et al. Astrophys. J. Suppl.* (in the press).
- Willmore, A. P. *et al. Mon. Not. R. astr. Soc.* 169, 7–23 (1974).
- Wilson, A. M., Carpenter, G. F., Eyles, C. J., Skinner, G. K. & Willmore, A. P. *Astrophys. J. Lett.* 215, L111–L115 (1977).
- Murdin, P. *et al. Mon. Not. R. astr. Soc.* 169, 25–34 (1974).
- Liller, W. *Astrophys. J. Lett.* 213, L21–L23 (1977).
- Swank, J. H., Becker, R. H., Pravdo, S. H. & Serlemitsos, P. J. *IAU Circ. No.* 2963 (1976).
- Li, F. K. *et al. Mon. Not. R. astr. Soc.* 179, 21p–25p (1977).
- Becker, R. H. *et al. Astrophys. J.* (in the press).
- Lewin, W. H. G., Hoffman, J. A. & Doty, J. *IAU Circ. No.* 3087 (1977).
- Seward, F. D., Page, C. G., Turner, M. J. L. & Pounds, K. A. *Mon. Not. R. astr. Soc.* 175, 39p–46p (1976).
- Lewin, W. H. G. *et al. Nature* 267, 28–30 (1977).
- Giacconi, R. *et al. Astrophys. J. Suppl.* 27, 37–64 (1974).
- Clark, G. W., Markert, T. H. & Li, F. K. *Astrophys. J. Lett.* 199, L93–L96 (1975).

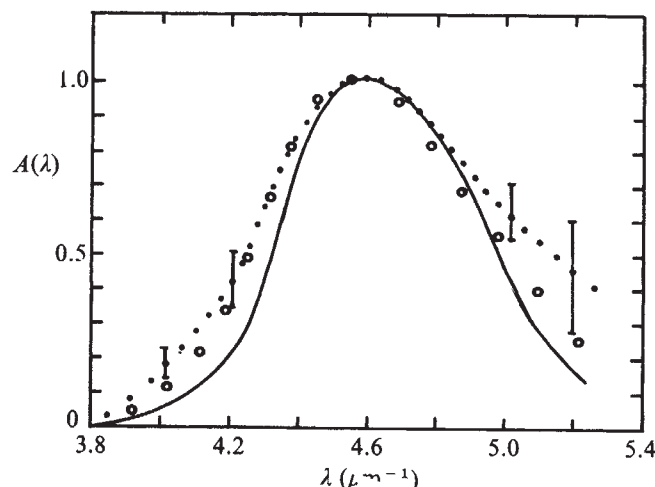


Fig. 1 Normalised average molar absorptivity for $C_8H_6N_2$ isomers (solid curve) compared with interstellar extinction data in the waveband $3.8 \mu m^{-1} < \lambda^{-1} < 5.4 \mu m^{-1}$. Normalisation is to 0.0 at $\lambda^{-1} = 3.8 \mu m^{-1}$, 1.0 at $\lambda^{-1} = 4.55 \mu m^{-1}$. Vertical bars give indication of spread of astronomical data. Dotted curve is the mean extinction curve of Bless and Savage² normalised according to equation (1). Open circles give mean extinction $(E(\lambda - V)/E(B - V))$ relative to extinction data for θ -Orionis, and normalised as above.

underlying graphite particulate extinction (scattering + absorption) peak upon which the narrower molecular absorption band is superposed. Since θ -Orionis shows a broad extinction hump centred on $\lambda 2,200\text{\AA}$ rather than the sharper band which is more common, we can reasonably attribute this extinction curve to an underlying graphite component. The mean extinction curve relative to the extinction data for θ -Orionis (Fig. 1) provides much closer agreement with the molecular absorption data, as we would expect. The mass density of molecules necessary to produce the observed strength of the $\lambda 2,200\text{\AA}$ interstellar band ($\sim 1.5 \text{ mag kpc}^{-1}$) is $\sim 10^{-27} \text{ g cm}^{-3}$ implying that only $\sim 10\%$ of interstellar C and N is in this form.

An identification of ring compounds of the type listed in Table 1 may have important consequences. Linear molecules such as HCN, HC_3N , HC_5N , HC_7N which have already been observed in interstellar space may result from the break-up of

Table 1 Properties of various isomers of $C_8H_6N_2$

Compound	Structural formula	$\lambda_m(\text{\AA})$	ϵ (molar absorptivity)
Cinnoline		2,250	40,000
Quinazoline		2,220	35,500
Quinoxaline		2,340	23,400
Phthalazine		2,150	56,000
1,5 Naphthyridine		2,060	54,000

these more complex structures. It would now be worthwhile to search systematically for ring molecules by radioastronomical techniques.

Sakata *et al.*⁸ have reported the detection of an absorption feature at $\lambda 2,200\text{\AA}$ in soluble organic material extracted from the Murchison meteorite. In view of the possible connection of this material with interstellar matter, a chemical analysis of the

molecules responsible for the meteoritic $2,200\text{\AA}$ band will also be valuable. It is interesting that several nitrogen heterocyclic compounds, including purines, pyrimidines and pyrroles have recently been identified in carbonaceous chondrites⁹⁻¹¹.

F. HOYLE*

N. C. WICKRAMASINGHE

Department of Applied Mathematics and Astronomy,
University College,
Cardiff, UK

Received 19 August; accepted 1 October 1977.

*Honorary Professorial Fellow: Present address: Cockley Moor, Dockray, Penrith, Cumbria, UK.

1. Stecher, T. P. *Astrophys. J.* **157**, L125 (1969).
2. Bless, R. C. & Savage, D. *Astrophys. J.* **171**, 193 (1972).
3. Witt, A. N. & Lillie, C. F. in *Scientific Results from the Orbiting Astronomical Observatory OAO-2* (ed. Code, A. (NASA SP-310, 1972)).
4. Hoyle, F. & Wickramasinghe, N. C. *Nature* (in the press).
5. Gilra, D. P. in *Scientific Results from the Orbiting Astronomical Observatory OAO-2* (ed. Code, A.) (NASA SP-310) (1972).
6. Hoyle, F. & Wickramasinghe, N. C. *Nature* **266**, 241-242 (1977).
7. *UV Atlas of Organic Compounds* (Butterworths, London, 1966).
8. Sakata, A. *et al.* *Nature* **266**, 241 (1977).
9. Hayatsu, R. *Science* **146**, 1291 (1964).
10. Hayatsu, R., Studier, M. H., Oda, A., Fuse, K. & Anders, E. *Geochim. cosmochim. Acta* **32**, 175 (1968).
11. Folsome, C. E., Lawless, J. G., Romiez, M. & Ponnampuram, C. *Geochim. cosmochim. Acta* **37**, 455 (1973).

Has the Sun a companion star?

PULSARS are accurate timekeepers. They are believed to be rotating neutron stars, with strong magnetic fields, and the energy they radiate is at the expense of their rotational kinetic energy¹. As each pulsar ages, its period P (relative to the Solar System barycentre) slowly increases, and its period derivative $\dot{P}$ ($= dP/dt$) slowly decreases. Certain interesting pulsars (displayed in Table 1) have anomalously small period derivatives, and rather surprisingly, are found grouped together in the same region of the sky (shown in Fig. 1). I suggest here, as an explanation of the peculiar properties of these pulsars, that the barycentre of the Solar System is accelerated, possibly because the Sun is a member of a binary system and has a hitherto undetected companion star.

The characteristic spindown age of any pulsar is $T = P/\dot{P}$. Most pulsars are born close to the galactic plane and most pulsars have high velocities; hence, as they age, they disperse away from the galactic plane. By studying the observed height distribution of pulsars about the galactic plane, various authors¹⁻⁵ have drawn the following conclusions.

First, the 'young' pulsars of $T \leq 3 \times 10^6$ yr (or $\dot{P} \geq 10^{-14}$) have an observed height distribution that shows a rise in scale-height with increasing T , that is consistent with the known velocities, and the spindown age is a reasonable approximation of true age.

Second, most pulsars (57 out of 85) of known period derivatives^{5,6} are 'middle-aged' and have spindown ages in the range $3 \times 10^6 \leq T \leq 3 \times 10^8$ yr (or $10^{-14} \geq \dot{P} \geq 10^{-16}$). In their case T is generally an unreliable indicator of true age, and at best is an upper limit. This is possibly because of magnetic-field decay^{1,2,8}, or change in magnetic-field orientation⁸⁻¹⁰, or perhaps because of other unknown processes. The positions of these 57 pulsars in equatorial coordinates are shown in Fig. 1; they are distributed about the galactic equator and show concentration toward the galactic centre.

Third, the remaining five pulsars of $T \geq 3 \times 10^8$ yr, listed in Table 1, have anomalously low period derivatives of $\dot{P} < 10^{-16}$. PSR1913+16 is the binary pulsar discovered by Hulse and Taylor¹¹, and its low period derivative $\dot{P} = 9 \times 10^{-18}$ may be due to exchange of angular momentum. There are, at present, no explanations for the anomalously low period derivatives of the remaining four pulsars.

The period derivatives shown in Table 1 must be further reduced by the Shklovsky¹⁶ or 'train-whistle' effect (whistle-blowing trains passing in all directions through a nearby railway station have positive $\dot{P}$). If $\dot{P}_0$ is the intrinsic period derivative of a source

moving with transverse velocity v_t at distance d , and $\dot{P}$ is the observed period derivative, then

$$\frac{\dot{P}}{P} = \frac{\dot{P}_0}{P} + \frac{v_t^2}{cd} \quad (1)$$

where c is the velocity of light. For PSR1944+17 the Shklovsky effect reduces $\dot{P}_0$ to almost zero¹⁷; and for PSR1952+29 it is found that $\dot{P}_0$ is negative¹⁵ and the pulsar is apparently spinning up. Mansfield and Rankin¹⁵ have investigated various mechanisms that might explain the negative intrinsic period derivative of PSR1952+29, and conclude that all hitherto proposed mechanisms are either inadequate or unlikely.

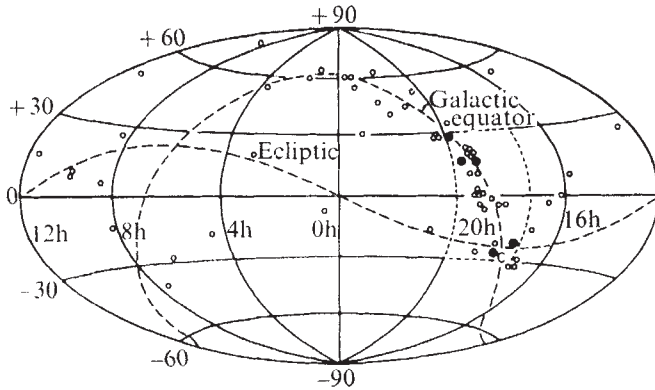


Fig. 1 Plot of pulsar positions in equatorial coordinates on an equal area projection. O, Pulsars of $10^{-14} > \dot{P} \geq 10^{-16}$; ●, pulsars of $\dot{P} < 10^{-16}$; C, Galactic centre.

The pulsars of small period derivative in Table 1 are shown with their positions plotted in Fig. 1. It is remarkable that all five lie in the same region of the sky with their right ascension (α) and declination (δ) within the ranges $17^{\text{h}}30' < \alpha < 19^{\text{h}}52'$ and $-27^\circ < \delta < +32^\circ$. Inside the box (shown in Fig. 1), defined by these coordinate limits, are found 23 pulsars of which five have the anomalously small $\dot{P}$; outside the box are the remaining 39 pulsars, all of which have $\dot{P} \geq 10^{-16}$. The probability of a chance clustering of five pulsars in 23 out of 62 is about 1%. Even if chance explains the clustering effect, it does not explain the anomalously low $\dot{P}$ values.

The above results suggest that the barycentre of the Solar System is accelerated in a direction that is within or close to the coordinate values already given. If $\dot{V}$ is the acceleration of the Solar System in a direction defined by coordinates α_0 , δ_0 , and $\Delta\dot{P}$ is the change in $\dot{P}$ of a pulsar located at coordinates α , δ , then

$$\frac{\Delta\dot{P}}{P} = -\frac{\dot{V}}{c} \{ \sin\delta \sin\delta_0 + \cos\delta \cos\delta_0 \cos(\alpha - \alpha_0) \} \quad (2)$$

Evidently, $-\Delta\dot{P} \sim 10^{-16}$ is sufficient to account for the low values of $\dot{P}$ in Table 1, including correction for the Shklovsky effect, and because $P \sim 1$ s, this corresponds to an acceleration $\dot{V} \sim 10^{-6} \text{ cm s}^{-2}$. This acceleration is very roughly in the direction of the galactic centre but is at least two orders of magnitude greater than the acceleration of the Sun in the gravitational field of the Galaxy. (More precise values of $\dot{V}$, α_0 , δ_0 cannot be found by averaging a weighted $\dot{P}$ over the celestial sphere. This is because of a selection effect: the pulsars in the direction of the galactic centre include those of higher displacement from the galactic plane and therefore they have lower average $\dot{P}$.)

An intriguing explanation of the acceleration $\dot{V}$ is that the Sun has an orbiting companion star. The acceleration is then

$$\dot{V} = GM_c/R_c^2 = 0.6 M/R^2 \text{ cm s}^{-2} \quad (3)$$

and is directed toward the companion star of mass M_c at distance R_c , where $M (= M_c/M_\odot)$ is its mass measured in units of the Sun's

Table 1 Pulsars of $\dot{P} < 10^{-16}$

Pulsar	P (s)	$\dot{P}$ (10^{-16})	Age (10^8 yr)	Distance (pc)	Refs
1730-22	0.87	0.0 ± 0.2	> 10	1,600	2,5
1813-26	0.59	-3 ± 3	?	3,500	2
(1913+16)	0.06	0.088 ± 0.003	~ 3	6,200	12)*
1944+17	0.44	0.244 ± 0.005	6	550	13
1952+29	0.43	0.02 ± 0.001	70	265	14,15

*Binary pulsar

mass and R its distance in AU. A companion star of mass $M \sim 1$ at distance $R \sim 10^3$ therefore produces an acceleration that is approximately equal to $\dot{V} \sim 10^{-6} \text{ cm s}^{-2}$. The circular orbital period of the hypothetical star is

$$P_c = R^{3/2}(1+M)^{-1/2} \text{ yr} \sim 10^4 \text{ yr} \quad (4)$$

and the corresponding proper motion is ~ 100 arc s per yr (the annual parallax is ~ 200 arc s).

Most nearby stars are members of systems that contain two or more stars¹⁸, and several stars have unseen companions detected astrometrically. The possibility of the Sun possessing a faint companion star has so far not been entirely eliminated¹⁹. A hypothetical companion star of, say, an apparent magnitude $m_v = 10$ (much too faint to be seen by the naked eye), has a luminosity $\sim 10^{-9}$ times that of the Sun. It is conceivable that the companion star is a crystallised white dwarf^{20,21} that has cooled rather rapidly to a surface temperature of less than 10^3 K during the lifetime of the Solar System. A red (or even black) dwarf²² of mass $M \lesssim 0.1$, that has failed to ignite hydrogen (but may be burning deuterium), has an exceedingly low luminosity and is therefore also a plausible candidate. If, however, the companion is a neutron star or black hole, produced by a supernova in the early history of the Solar System, then it is difficult to understand how our weakly bound binary system has survived such a violent event. It is more likely that a companion neutron star or black hole is in hyperbolic orbit, and the present close encounter is a transitory phenomenon lasting only several thousands of years. The companion star is therefore presumably either a faint white or red dwarf in closed orbit about the Sun, or a gas-accreting nearby neutron star or black hole in open orbit. It should not be difficult to detect low-luminosity objects of this kind having large annual parallax and proper motion.

Obviously, a star in closed orbit about the Sun must lie close to the ecliptic plane so as not to disturb the planets excessively out of their common orbital plane. According to current theories on star formation²³, stars and their planets condense in interstellar gas clouds, and therefore planets probably revolve in the same orbital plane as the parent stars of a binary system. A companion star orbiting close to the ecliptic is not inconsistent with the distribution of pulsars of $\dot{P} < 10^{-16}$ shown in Fig. 1.

A companion star in closed orbit would produce various long-term effects within the Solar System. For example, the apsidal period U (the period of precession of perihelion) of a planetary orbit of period P_p is^{24,25}

$$U = \frac{4}{3}(1+M^{-1})P_c^2P_p^{-1} \quad (5)$$

and hence, with $M \sim 1$, $P_p \sim 1$ yr, $P_c \sim 10^4$ yr, this equation yields $U \sim 10^8$ yr, and the precession amounts 0.5 arc s per century (or 0.1 arc s per century in the case of Mercury). Short-term perturbations in the motions of Neptune and Pluto are significant, but as far as I can determine, are not unacceptably large.

Long-term changes in climate are conceivably attributable to periodicities in the Earth's orbit^{26,27}, that could be enhanced by the presence of a companion star, particularly if its orbit has appreciable ellipticity. Furthermore, cometary motions will undoubtedly be affected. According to Oort's comet-cloud theory²⁸, the comets extend out to a distance $\sim 10^5$ AU, and the comet cloud

therefore envelops the proposed binary system of the Sun and its companion star. Oort has suggested that the nearby stars, at a few light years distance, perturb the comet cloud and as a result comets continually diffuse into orbits of comparatively small perihelia. The proposed binary system would obviously play an even more effective role in stirring the cometary orbits, and could also explain why the comet cloud, during the lifetime of the Solar System, has expanded to its present size.

Has the Sun a companion star? I find it hard to believe that a star so close can exist and yet remain undiscovered. On the other hand, pulsar observations of extraordinary precision imply that it might exist, and therefore a search for a companion star is perhaps worth undertaking.

I thank T. T. Arny, D. J. Helfand, W. C. Saslaw, and particularly J. H. Taylor, for helpful comments, and also to the Aspen Center for Astrophysics for hospitality. Helfand has recently measured $\dot{P}$ for PSR2106+44 and finds that for this pulsar: $\dot{P} = -1.36 (\pm 1.00) \times 10^{-17}$. It is interesting that this pulsar is in the same region of the sky as the other pulsars of anomalously low $\dot{P}$. I also thank Dr P. van de Kamp for his comments, and for pointing out that the 'train whistle effect' has been known since 1901 as 'perspective secular change in the line of sight'²⁹.

E. R. HARRISON

Department of Physics and Astronomy,
University of Massachusetts,
Amherst, Massachusetts 01003

Received 2 August; accepted 13 October 1977.

- Gunn, J. E. & Ostriker, J. P. *Astrophys. J.* **160**, 979 (1970).
- Lyne, A. G., Rutchings, R. T. & Smith, F. G. *Mon. Not. R. astr. Soc.* **171**, 579 (1975).
- Harrison, E. R. & Tademaru, E. *Astrophys. J.* **201**, 447 (1975).
- Helfand, D. J. & Tademaru, E. *Nature* **267**, 130 (1977).
- Helfand, D. J. & Tademaru, E. *Astrophys. J.* **216**, 842 (1977).
- Taylor, J. H. & Manchester, R. N. *Astr. J.* **80**, 794 (1975).
- Terzian, Y. & Davidson, K. *Astrophys. Space Sci.* **44**, 479 (1976).
- Flowers, E. & Ruderman, M. A. *Astrophys. J.* **215**, 302 (1977).
- Macy, W. W. *Astrophys. J.* **190**, 153 (1974).
- Jones, P. B. *Nature* **262**, 120 (1976).
- Hulse, R. A. & Taylor, J. H. *Astrophys. J. Lett.* **195**, L51 (1975).
- Taylor, J. H., Hulse, R. A., Fowler, L. A., Gullahorn, G. E. & Rankin, J. M. *Astrophys. J. Lett.* **206**, L53 (1976).
- Richards, D. W., Rankin, J. M. & Zeissig, G. A. *Nature* **251**, 37 (1974).
- Gullahorn, G. E., Payne, R. R., Rankin, J. M. & Richards, D. W. *Astrophys. J. Lett.* **205**, L151 (1976).
- Mansfield, V. N. & Rankin, J. M. *Vistas in Astronomy* (in the press).
- Shklovsky, I. S. *Astrophys. Lett.* **8**, 101 (1970).
- Helfand, D. J., Taylor, J. H. & Manchester, R. N. *Astrophys. J. Lett.* **213**, L1 (1977).
- Van de Kamp, P. A. *Rev. astr. Soc.* **9**, 103 (1971).
- Van de Kamp, P. A. *Pub. astr. Soc. Pac.* **73**, 404 (1961).
- Van Horn, H. M. *IAU Symp. White Dwarfs* (Reidel, Dordrecht, 1970).
- Ostriker, J. P. A. *Rev. astr. Soc.* **9**, 353 (1971).
- Kumar, S. S. *Low-Luminosity Stars 255* (Gordon and Breach, New York, 1969).
- Jastrow, R. & Cameron, A. G. W. (eds) *Origin of the Solar System* (Academic, New York, 1963).
- Slavenas, P. *Trans. astr. Obs. Yale Univ.* **6**, 35 (1927).
- Kopal, Z. *Close Binary Systems* 108 (Wiley, New York, 1959).
- Ward, W. R. *J. geophys. Res.* **79**, 3375 (1974).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121 (1976).
- Oort, J. H. *Bull. astr. Inst. Neth.* **11**, 259 (1951).
- van de Kamp, P. *Elements of Astromechanics* 128 (Freeman, San Francisco, 1964).

Lines of H₂ in extreme-ultraviolet solar spectra

THE first detection of H₂ in the Sun, in extreme ultraviolet spectra, is reported here. The Naval Research Laboratory's High Resolution Telescope and Spectrograph (HRTS), flown in a rocket on 21 July 1975, was used for these observations. The slit extended from Sun centre to limb and crossed a sunspot at 07° S, 23° W in McMath region 13766. This stigmatic spectrograph, described by Bartoe and Brueckner¹, covered 1,175–1,714 Å with 0.06 Å spectral and 1.5 arc s spatial resolution. A full list of observed lines is in preparation². The sunspot spectrum contains some 120 emission lines which do not correspond to known lines listed in compilations³ of transitions in atoms or ions, apart from random coincidences. More recent publications referred to in the NBS Bibliography⁴ were also consulted. The identifications of the lines which are enhanced in the sunspot spectrum will be discussed elsewhere⁵.

A short stretch of a spectrum obtained with a 51-s exposure time is reproduced in Fig. 1. The top spectrum is of the solar limb,

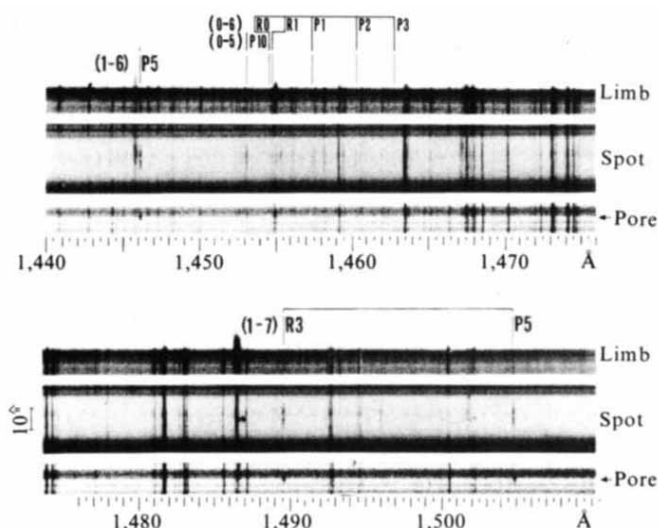


Fig. 1 A short section of the stigmatic ultraviolet solar spectrum showing several of the Lyman band lines in the 1,440–1,510 Å region. Each of the two sections covers three different areas on the solar limb spectrum is at the top, the spot spectrum in the centre and the 'pore' spectrum at the bottom.

the centre one is of the sunspot and the lower one is of a region of the quiet Sun which contains a small area (~ 2 arc s) in which some of the sunspot lines are greatly enhanced (a 'pore'?).

At least three distinct types of lines are present. Some are observed in the spot, pore and at the limb (for example, the line at 1,446.13 Å in Fig. 1); others are restricted to the spot and limb or to the pore and limb.

The continuum in both the sunspot umbra and penumbra is below the instrumental threshold for detection even in the longest exposure of 51 s. At 1,600 Å this corresponds to an upper limit of 3,950 K for the region of formation of the continuum, compared with the normal quiet Sun value of $\sim 4,400$ K. From their visible region spectra sunspots are found to have photospheric temperatures of $\sim 4,400$ K, and the temperature minimum above a spot is typically considered to be $\sim 3,200$ K (ref. 6). The visible region spectra are therefore characterised by the presence of low excitation absorption lines and by hundreds of absorption lines of molecular origin which do not appear in the normal photospheric spectrum. In the ultraviolet the known emission lines of singly ionised species decrease in intensity over the spot and only neutral lines formed deepest in the chromosphere are enhanced. The origin of the new lines is therefore restricted to species with low ionisation potentials or dissociation energies.

Because of the high abundance of hydrogen and the expected low temperatures, the molecule H₂ (dissociation energy 4.55 eV), which has strong far ultraviolet bands, is an obvious candidate for any new lines in the sunspot. An early comparison with known laboratory spectra⁷ of H₂, however, indicated no general agreement in the character of the spectra. In a 100 Å section near 1,500 Å there are of the order of 100 vibrational lines in the Lyman bands, whereas the sunspot spectrum over the same wavelength range shows only a few lines. Moreover, even allowing for different excitation conditions between the Sun and laboratory, most of the lines expected to be strong are absent.

But it has now been found that at least 30 of the new lines are due to known transitions in the Lyman bands of H₂, the key to the appearance of the spectrum lying in the excitation mechanism for the lines.

The Lyman bands arise between the ground state $1\Sigma_g^+$ and the excited state $1\Sigma_u^+$. The lines of many vibrational bands have been tabulated by Herzberg and Howe⁷, and a useful *Atlas of Intense Lines* has been prepared by Schubert and Hudson (unpublished Aerospace Report, ATN-64(9233)-2).

The most striking aspect of the lines which correspond to H₂ is that 13 of them can be identified at P-branch ($J = 5$) and R-branch

($J=3$) transitions in the $v'=1$ to $v''=n(2 < n < 9)$ Lyman bands; only one line in each branch is observed in each band, all originating from a common $v'=1, J'=4$ upper level. Progressions with upper vibrational quantum number $v'=constant$ were first studied by Wood and Hackett^{8,9} and are usually due to fluorescence.

The source of the fluorescence is not hard to find, given the high intensity of the nearby H L α line. The energy difference between the $v'=1, J'=4$ level in $^1\Sigma_u^+$ and the $v''=2, J''=5$ level in $^1\Sigma_g^+$ is within 26 cm^{-1} of the energy of H L α . The full-width at half maximum (FWHM) of H L α is $\sim 1\text{ \AA}$. So the H₂ molecules in the sunspot, or pore, can be excited by photons in the red wing of H L α as they travel down towards the photosphere.

The observed lines from $v'=1$ are listed in Table 1, together with their intensities in the umbra, and their FWHM. The theoretical intensity ratio¹⁰ of P5 to R3 lines is 10/8, and within the limits of experimental accuracy the observed line intensities agree with this. Along the progression the intensities follow closely the relative strengths of the bands as predicted by the transition probabilities of Allison and Dalgarno¹¹. The observed line widths correspond to a 'thermal' temperature of $7.7 \times 10^3\text{ K}$, or with $T_e \sim 4,400\text{ K}$, a nonthermal velocity of $\sim 5\text{ km s}^{-1}$.

Table 1 Lines observed in the $v'=1$ to $v''=n(2 < n < 9)$ Lyman bands of H₂

Transition $v'v''$	λ^* ($\text{\AA}$)	Intensity ($\text{erg cm}^{-2}\text{s}^{-1}\text{st}^{-1}$)	FWHM ($\text{\AA}$)	Comments
1-2 P5	(1,216.05)	—	—	Masked and excited by L α
R3	1,202.45	—	—	Weak in umbra
1-3 P5	1,271.93	11	0.090	
R3	1,257.80	11	0.083	
1-4 P5	(1,329.14)	—	—	Masked by C(i) 1,329.10
R3	1,314.62	3	—	
1-5 P5	1,387.35	2	—	
R3	1,372.49	—	—	Weak in umbra
1-6 P5	1,446.13	9.6	0.093	
R3	1,431.01	7.4	0.088	
1-7 P5	1,504.79	14	0.093	
R3	1,489.57	12	0.094	
1-8 P5	1,562.41	12	0.114	
R3	1,547.35	14	0.101	
1-9 P5	1,617.93	—	—	Weak in umbra
R3	1,603.24	—	—	Weak in umbra

*The wavelengths are from Herzberg and Howe⁷. Solar wavelengths agree with laboratory wavelengths within 0.02 $\text{\AA}$.

The 1-0 and 1-1 bands are outside the wavelength range of the instrument, but an 'anti-Stokes' line is observed at 1,202.45 $\text{\AA}$. Other fluorescence occurs, accounting for transitions in the 0-4, 0-5 and 0-6 bands. To produce these the $v'=0, J'=1$ and 2 levels are excited from $v''=2, J''=1$ by photons 104 cm^{-1} and 134 cm^{-1} out into the red wing of H L α . The $J''=9$ level is also excited but not apparently by photons. But its energy is close to that of the $v'=1, J'=4$ level discussed above and collisional processes may be indicated. This second group of lines is listed in Table 2. There are a further 20 or so weak lines of similar appearance which correspond by wavelength to H₂ transitions but which do not show such systematic behaviour.

The closest coincidence with H L α , the transition between $v'=3, J'=1$ and $v''=3, J''=0$, does not produce observable lines. This is probably due to the combined effects of the self-reversal of H L α , a lower transition probability and a lower population of $v''=3$.

The remaining lines, which do not appear in the pore, may also have a molecular origin, but the singly ionised rare-earths are another possible source of low excitation lines.

The aspects of the data which relate to the understanding of the solar atmosphere will be discussed in the forthcoming paper⁵. But

Table 2 Lines observed in the $v'=0$ to $v''=4,5,6$ Lyman bands of H₂

Transition $v'v''$	λ^* ($\text{\AA}$)	Intensity ($\text{erg cm}^{-2}\text{s}^{-1}\text{st}^{-1}$)	FWHM ($\text{\AA}$)	Comments
0-4 P1	(1,335.87)	—	—	Masked by C(ii) 1,335.70
P2	1,338.57	7.2	0.083	
P3	1,342.26	—	—	Weak
P10	1,393.45	4.0	—	Blend?
R0	1,333.48	5.4	0.099	
R1	1,333.80	—	—	Weak
R8	1,363.58	5.1	0.092	
0-5 P1	1,396.22	—	—	Weak
P2	1,398.96	—	—	Weak
P3	(1,402.65)	—	—	Masked by Si(IV) 1,402.79
P10	1,453.02	7.7	0.105	
R0	(1,393.72)	—	—	Masked by Si(IV) 1,393.78
R1	(1,393.96)	—	—	Masked by Si(IV) 1,393.78
R8	1,422.55	1.8	0.085	
0-6 P1	1,457.43	—	—	Weak
P2	1,460.17	2.4	0.068	
P3	1,463.83	2.6	0.095	
P10	—	—	—	
R0	1,454.83	—	—	Weak
R1	1,454.97	3.2	0.091	b1 1455.00?
R8	—	—	—	

*The wavelengths are from Herzberg and Howe. Solar wavelengths agree with laboratory wavelengths within 0.02 $\text{\AA}$.

the importance of this H L α , H₂ fluorescence process extends beyond the solar atmosphere, since it will occur wherever strong L α radiation can penetrate a layer of temperature $\sim 2,000\text{--}5,000\text{ K}$. For example, one can immediately predict that these H₂ lines should be present in stars cooler than the Sun, particularly those of type early K, since the visible region spot 'special type' is K0. Such observations could be made with the forthcoming International Ultraviolet Explorer (IUE) Satellite, which includes the appropriate wavelength range. Further, Osterbrock¹² and Black and Dalgarno¹³ have stressed the importance of the ultraviolet radiation field in the formation of the interstellar lines of H₂. The penetration of L α from hot stars into nearby cooler media may play a particularly important role, although the low population of the $v''=2$ levels will make fluorescence less effective than in stellar atmospheres.

C. JORDAN

Department of Theoretical Physics,
University of Oxford,
1 Keble Road, Oxford, UK

G. E. BRUECKNER
J.-D. F. BARTOE
G. D. SANDLIN
M. E. VAN HOOSIER

E. O. Hulbert Center for Space Research,
Naval Research Laboratory,
Washington, D.C., 20375

Received 17 September; accepted 11 October 1977.

1. Bartoe, J.-D. F. & Brueckner, G. E. *J. opt. Soc. Am.* **65**, 13, 21 (1975).
2. Sandlin, G. D., Bartoe, J.-D. F., Brueckner, G. E., Van Hoosier, M. E., *Astrophys. J. Suppl. Ser.* (to be submitted).
3. Kelly, R. L. & Palumbo, L. J. *NRL Rep.* 7599 (1973).
4. Hagan, L. *NBS Spec. Publ.*, 363, Suppl. 1 (1977).
5. Jordan, C., Brueckner, G. E., Bartoe, J.-D. F., Sandlin, G. D. & Van Hoosier, M. E., *Astrophys. J.* (to be submitted).
6. Allen, C. W. *Astrophysical Quantities* 3rd edn, 185 (Athlone, London, 1973).
7. Herzberg, G. & Howe, L. L. *Can. J. Phys.* **37**, 636, 659 (1959).
8. Wood, R. W. *Phil. Mag.* **12**, 499-522 (1906).
9. Wood, R. W. & Hackett, F. E. *Astrophys. J.* **30**, 330-372 (1909).
10. Herzberg, G. *Spectra of Diatomic Molecules* (Van Nostrand, Princeton, 1950).
11. Allison, A. C. & Dalgarno, A. *Atomic Data* **1**, 289-304 (1970).
12. Osterbrock, D. E. *Astrophys. J.* **136**, 359-364 (1962).
13. Black, J. H. & Dalgarno, A. *Bull. Am. astr. Soc.* **6**, 444 (1974).

Significance of peroxyntiric acid in atmospheric chemistry of nitrogen oxides

KNOWLEDGE of the life cycle and budget of nitrogen oxides in the atmosphere is important for several reasons. In the stratosphere, NO and NO₂ (NO_x) play a dominant part in maintaining the structure and stability of the ozone distribution. In the troposphere, NO_x is involved in the production and destruction of OH radicals which control the fluxes of many important species such as CH₄, CO, H₂ and CH₃Cl, to the stratosphere. In polluted urban areas, NO_x chemistry provides the basic formation mechanism of photochemical air pollution. In all these areas of atmospheric chemistry, the interactions of NO_x with HO_x species (HO, HO₂ and so on) are of great significance, particularly the reactions of HO₂ with NO and of HO with NO₂. A further HO_x-NO_x interaction is the proposed¹ formation of peroxyntiric acid in reactions (1) and (-1):



We present here a study of the extent to which reaction (1) and (-1) acts as a sink for HO_x and NO_x in the atmosphere.

Peroxyntiric acid was first reported to possess rather low stability at room temperature in the gas phase¹, decomposing back to HO₂ and NO₂ with a lifetime of about 50 s. Subsequently other groups^{2,3} have identified infrared absorptions of HO₂NO₂, reporting much longer lifetimes, about 500 s, with the decay being ascribed to heterogeneous removal at the vessel walls. These apparently diverging conclusions can be rationalised if HO₂NO₂ is in thermal equilibrium with HO₂ and NO₂ as we have shown to be the case for its organic analogue PAN (peroxyacetyl nitrate) and the corresponding peroxy radical, CH₃CO(O₂) (ref. 4). When NO₂ is in large excess over NO, as in the infrared experiments, the main fate of the HO₂ radical is to recombine with NO₂ reforming HO₂NO₂ again. Only when nitric oxide can effectively compete against NO₂ for HO₂, will HO₂NO₂ decay at its unperturbed unimolecular decay rate, k_{-1} .

We have studied the temperature dependence of this competition for HO₂ radicals in reactions (1) and (5) by producing HO₂

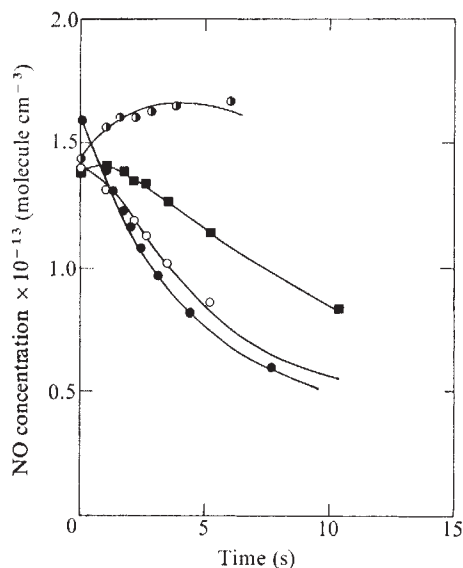
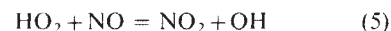
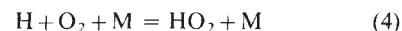
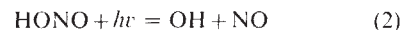


Fig. 1 NO concentration-time curves showing the effect of increasing NO₂ on the inhibition of NO oxidation in mixtures containing approximately 2×10^{14} and 5×10^{16} molecule cm⁻³ of HONO and CO in 1 atm of air at 35 °C. Initial NO_x concentrations, [NO₂]₀ in molecule cm⁻³ are ●, 1.78×10^{13} ; ○, 3.24×10^{13} ; ■, 5.90×10^{13} ; ○, 1.15×10^{14} . Lines show computed curves used in estimating k_1/k_5 , and k_{-1} .

radicals from the photolysis of nitrous acid (HONO) in the presence of carbon monoxide (CO):



The photolysis was conducted in a flow system⁵ with reaction times of 1–13 s. NO_x was present at p.p.m. concentrations in synthetic air diluent at 1 atm pressure; NO, NO₂ and HONO were measured by chemiluminescence techniques⁶.

The observed time dependence of the nitric oxide concentration, [NO], reflects the balance between production by photolysis of HONO and NO₂ and by oxidation through the chain reaction sequence, (3) + (4) + (5) (see Figs 1 and 2). Figure 1 shows that [NO] usually declines, indicating the dominance of chain oxidation over photolytic production. But increasing the NO₂ concentration leads to decreased nitric oxide decay rates. The magnitude of the observed effect requires HO₂NO₂ formation to be the explanation

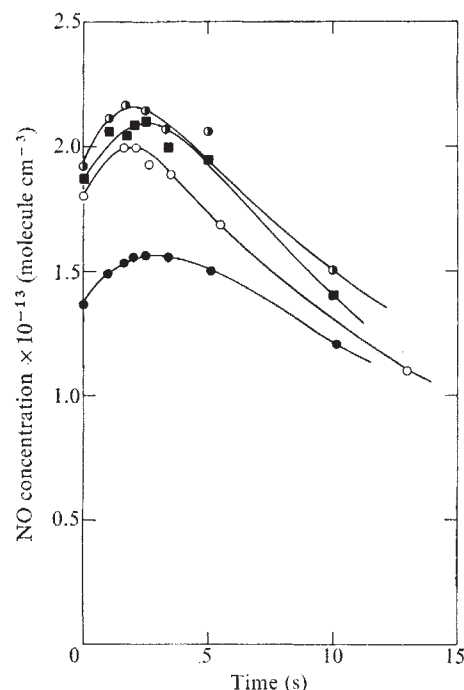


Fig. 2 NO concentration-time curves showing the effect of temperature on the inhibition of NO oxidation by the addition of approximately 1×10^{14} molecule cm⁻³ NO₂ to mixtures containing HONO and CO in 1 atm of air. Experimental points: ●, 40 °C; ■, 45 °C; ○, 49 °C; ○, 55 °C. Lines show computed curves using $k_1/k_5 = 0.22$ and k_{-1} values illustrated in Fig. 3.

rather than increased NO₂ photolysis. Figure 2 shows that the inhibiting effect of HO₂NO₂ formation is only transitory. The initial increase in [NO] due to strong inhibition of the chain oxidation gives way to a steady decline at a rate almost independent of temperature. The induction period, representing the time taken for [NO] to reach its maximum, increases with decreasing temperature. This behaviour is exactly that expected for the formation of an intermediate which is thermally unstable with respect to active free radical chain carriers. Thus peroxyntiric acid is apparently a kinetically significant species in the HO_x + NO_x system at low temperatures.

Quantitative kinetic information was obtained by integrating the differential equations describing the reaction scheme, using the Harwell computer program FACSIMILE⁷ and literature rate constants⁸. The values of k_1 and k_{-1} were obtained by minimising

the sum of the squares of the deviations between observed and simulated $[\text{NO}]$ against time curves. The experiments below 295 K gave consistent values of 0.22 ± 0.03 for the rate constant ratio k_1/k_5 in air at 1 atm total pressure. In these conditions the HO_2NO_2 lifetime was longer than the residence time in the flow reactor so that redissociation was unimportant and the $[\text{NO}]$ against time profile was determined primarily by the relative rates of reactions (1) and (5). This ratio k_1/k_5 was then used at the higher temperatures to obtain the values of k_{-1} which are plotted in Arrhenius form in Fig. 3. A value of $1.2 \times 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ (ref. 9) was adopted for k_5 but neither k_{-1} or k_1/k_5 were sensitive to the chosen value of k_5 in the range 10^{-11} to 10^{-12} . The values of k_{-1} are represented by $k_{-1} = 10^{16.1 \pm 1.6} \exp(-11,700 \pm 1,100/T) \text{ s}^{-1}$ for T in K at 1 atm pressure. The pre-exponential factor and activation energy, $(97 \pm 9) \text{ kJ mol}^{-1}$, are quite reasonable for the unimolecular decomposition of a weakly bound molecule into two radical fragments.

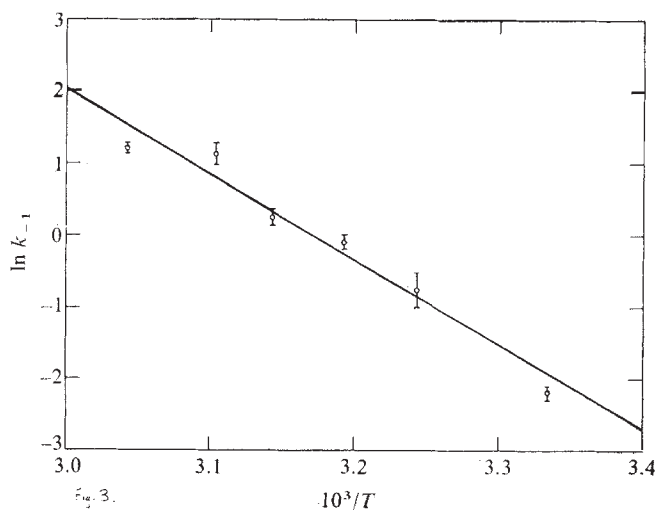


Fig. 3 Temperature dependence of the decomposition rate constant for HO_2NO_2 , k_{-1} ; plot of $\ln k_{-1}$ against $10^3/T$ (deg K^{-1}).

Decomposition lifetimes of HO_2NO_2 calculated from the above expression for k_{-1} give some indication of the stability of HO_2NO_2 at various altitudes in the atmosphere. In the coolest parts of the atmosphere near the tropopause, mean temperatures are about 220 K, corresponding to an HO_2NO_2 lifetime of about 16 weeks (0.3 yr) with respect to thermal decomposition. As a result, the concentration distribution of HO_2NO_2 in the upper troposphere and lower stratosphere will depend on the atmospheric circulation and dynamics as well as on the chemistry, provided that HO_2NO_2 is not removed rapidly by photolysis and reactions with other stratospheric species. Calculations with a two-dimensional model indicate that as much as 30% of the total NO_x in the upper troposphere may become tied up as HO_2NO_2 . This species may therefore play a similar role to HNO_3 in providing a removal mechanism for stratospheric NO_x .

In the atmosphere near the Earth's surface HO_2NO_2 is unstable with a decomposition lifetime of 9 s at 298 K. HO_2NO_2 will therefore not provide a significant sink or store for HO_2 or NO_2 , merely existing at a low steady-state concentration in the sunlit atmosphere. Consequently HO_2NO_2 is likely to have only a limited significance in photochemical air pollution and model calculations of photooxidant formation will not be seriously in error by omitting this species. This conclusion is in contrast to that of Levine *et al.*³ who suggest that HO_2NO_2 may be an important product of photochemical smog.

This work was sponsored by the Department of the Environment.

R. A. COX
R. G. DERWENT

Environmental and Medical Sciences Division,
AERE, Harwell, Didcot, Oxon, UK

A. J. L. HUTTON

Department of Chemistry,
University of Birmingham,
PO Box 363, Birmingham, UK

Received 15 August; accepted 3 October 1977.

1. Simonaitis, R. & Heicklen, J. *J. phys. Chem.* **80**, 1-7 (1976).
2. Niki, H., Maker, P. D., Savage, C. M. & Breitenbach, L. P. *Chem. Phys. Lett.* **45**, 564-566 (1977).
3. Levine, S. Z., Uselman, W. M., Chan, W. H., Calvert, J. G. & Shaw, J. H. *Chem. Phys. Lett.* **48**, 528-535 (1977).
4. Cox, R. A. & Rolley, M. J. *Environ. Sci. Technol.* **11**, 900-906 (1977).
5. Cox, R. A. *J. Photochem.* **3**, 175-188 (1974).
6. Atkins, D. H. F. & Cox, R. A. *A.E.R.E. Rep. R-7615* (Her Majesty's Stationery Office, London, 1973).
7. Curtis, A. R. & Kirby, C. R. *CSSD Rep. CSS 26; CSS 46* (Atomic Energy Research Establishment, Harwell, 1973).
8. Hampson, R. F. & Garvin, D. *NBS Technical Note 866* (US Department of Commerce, Washington, 1975).
9. Cox, R. A. & Derwent, R. G. *J. Photochem.* **4**, 139-153 (1975).

Vertical profile of stratospheric ozone by lidar sounding from the ground

LIDAR measurements of stratospheric ozone by differential absorption (DIAL) technique were taken during July 1977 at the Haute-Provence Observatory (44N). A frequency doubled rhodamine 6G dye laser provided an output energy of 20-40 mJ in the wavelength range 300-310 nm. We report here preliminary measurements of stratospheric ozone in the altitude range 18-25 km. This experiment was carried out with the lidar facility previously used for the study of upper atmospheric metallic constituents at the mesopause level¹.

The transmitting part of the lidar system consists of a flash-lamp pumped, frequency-doubled dye laser emitting in the wavelength range 300-310 nm. The experimental set-up is shown on Fig. 1 and its characteristics are summarised in Table 1. The dynamic range of this electronic system together with the close proximity of the emitting and receiving optical parts does not allow observations in a low altitude range (5-15 km); in order to avoid signal induced noise³ or spurious signal from photomultiplier overloading, a rotating wheel protects the photomultiplier during the first 100 μs . These constraints limit our experiments to an upper altitude range of 18-28 km. Observations have been conducted on two different nights (19-20 July, 20-21 July 1977) and more than 5×10^3 laser shots have been registered and analysed.

The signal backscattered by Rayleigh diffusion in the altitude range $z, z + \Delta z$ at a wavelength λ_1 can be written as

$$N_R(\lambda_1, z) = N_L(\lambda_1) \frac{A d\sigma_r}{z^2 d\Omega} (\lambda_1) n_a(z) \Delta z \eta T^2(\lambda_1, z) \exp[-2\tau_{03}(\lambda_1, 0, z)] \quad (1)$$

where $N_L(\lambda_1)$ is the number of emitted photons, A the telescope area, η the optical efficiency of the lidar system, $d\sigma_r/d\Omega(\lambda_1)$ the

Table 1 Characteristics of the lidar system

Transmitter		Receiver	
Energy per pulse	20-40 mJ	Telescope area	0.5 m ²
Output wavelength	303.7-308 nm	Beamwidth	5×10^{-3} rad
Linewidth	0.15 nm	Linewidth	3 nm
Beam divergence	$< 8 \times 10^{-4}$ rad		
Pulse duration	4 μs		
Pulse repetition rate	0.25 Hz		

The receiver is a 0.8-m Coudé telescope. Two interference filters were used for these observations: they are centred respectively at 308.2 nm and 303.7 nm with a peak transmission of 30% and a width of 3 nm (FWHM). These wavelengths have been selected to obtain a ratio of 1.75 in the absorption cross-sections of ozone and are compatible with an observation in the stratospheric range. The Rayleigh backscattered signal is detected by a trialkali cathode photomultiplier (EMR 541 E) and its temporal analysis is made by a multichannel analyser working in a photon counting mode with a vertical resolution of 1.2 km.

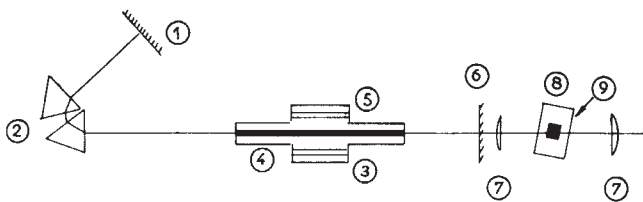


Fig. 1 The laser experimental set-up. 1, High reflectivity mirror; 2, wavelength selecting prisms; 3, laser head; 4, dye cell; 5, linear flash tubes; 6, output mirror ($R = 40\%$); 7, cylindrical lens; 8, crystal housing (UG5 output window); 9, ADP crystal. The oscillator head is made of three 15-cm length bielliptical cylindrical reflectors, successively mounted with a 60° angle between their plane of symmetry to obtain a total pumping length of 45 cm. Six linear flash tubes (VQX6FD15) are then used in this optical arrangement. The fundamental line (600–620 nm) is generated using an active solution of rhodamine 6G (10^{-4} Ml $^{-1}$) in water with 4% Ammonix LO. With water as a solvent, the length of the laser cavity has been increased up to 2 m to reduce the beam divergence to less than 8×10^{-4} radian, leading to a better efficiency in the Second Harmonic Generation (SHG)². Spectral narrowing of the emitted line to 0.15 nm is made using two quartz prisms working at their minimum deviation angle, which are also used to polarise the emitted beam. Rotation of the high reflectivity mirror leads to the tuning of the output wavelength which can be controlled using a grating spectrograph (dispersion 0.3 nm mm $^{-1}$). The output energy of 1 J which is usually obtained with an efficiency of 0.25% in a shot by shot experiment, has been reduced to half of this value for the atmospheric observations in order to increase the lifetime of the electrical components and to reduce the photolysis of the lasing solution. An ammonium dihydrogen phosphate crystal oriented for critical phase matching angle at the selected output wavelength is used to perform the SHG with an energy efficiency of 5% to 10% depending on the output power of the fundamental beam. Due to the presence of a 3 mm-thick UG5 window at the output edge of the crystal housing, only the ultraviolet light is emitted in order to avoid contamination effect of the backscattered signal by the fundamental visible beam.

Rayleigh backscattering cross-section, $n_a(z)$ the total atmospheric density and $T(\lambda_1, z)$ the atmospheric transmission between the ground and altitude z in the absence of ozone absorption.

$$\tau_{03}(\lambda_1, 0, z) = \int_0^z \sigma_{03}(\lambda_1) n_{03}(h) dh$$

is then the optical thickness of ozone at wavelength λ_1 on a one-way path between the ground and altitude z . Its value is derived

from the experimentally measured $N_R(\lambda_1, z)$ assuming that ozone is the main absorber at wavelength λ_1 between z and $z + \Delta z$. Interference effects with other minor constituents (SO_2 , H_2O , aerosols) are important in the troposphere⁴ but their concentration are thought to be too low in the stratosphere range to produce any detectable attenuation. Absorption by molecular oxygen is also negligible at 300 nm where the cross-sections are of the order of 10^{-26} cm 2 (ref. 5) which leads to an optical thickness of 10^{-3} for a 1.2 km resolution range.

The effect of aerosols around 22 km (ref. 6) should also be considered assuming a maximum extinction coefficient of 4×10^{-8} cm $^{-1}$ at 300 nm and thus an optical thickness of 5×10^{-3} which is again one order of magnitude lower than the expected ozone absorption. These assumptions will be confirmed later by comparing signals obtained at two different wavelengths. The only contribution to the atmospheric transmission $T(\lambda_1, z)$ is then the total molecular extinction by Rayleigh scattering which can be written as $T(\lambda_1, z) = \exp[-\tau_a(\lambda_1, 0, z)]$ where τ_a is the optical thickness defined as:

$$\tau_a(\lambda_1, 0, z) = \int_0^z \frac{d\sigma_r(\lambda_1)}{d\Omega} n_a(h) dh$$

From the ratio:

$$R(\lambda_1, z, z + \Delta z) = \frac{N_R(\lambda_1, z)}{N_R(\lambda_1, z + \Delta z)} \frac{(z + \Delta z)^2}{z^2} \frac{n_a(z)}{n_a(z + \Delta z)} \times \exp - 2\tau_a(\lambda_1, z, z + \Delta z) \exp - 2\tau_{03}(\lambda_1, z, z + \Delta z) \quad (2)$$

and using an atmospheric density model for midlatitude summer conditions⁷ to derive the values of $[n_a(z)/n_a(z + \Delta z)] \exp - 2\tau_a(\lambda_1, z, z + \Delta z)$, we can now deduce the optical thickness of the ozone layer between z and $z + \Delta z$. Figure 2a and b show these values of τ_{03} between 18 km and 28 km with a resolution of 1.2 km, as measured at two different wavelengths ($\lambda_1 = 308.2$ nm, $\lambda_2 = 303.7$ nm). The deviation from a pure Rayleigh atmosphere shows up clearly. The integration time, 3 h during the night of 20 July, corresponds to 3×10^3 laser shots at

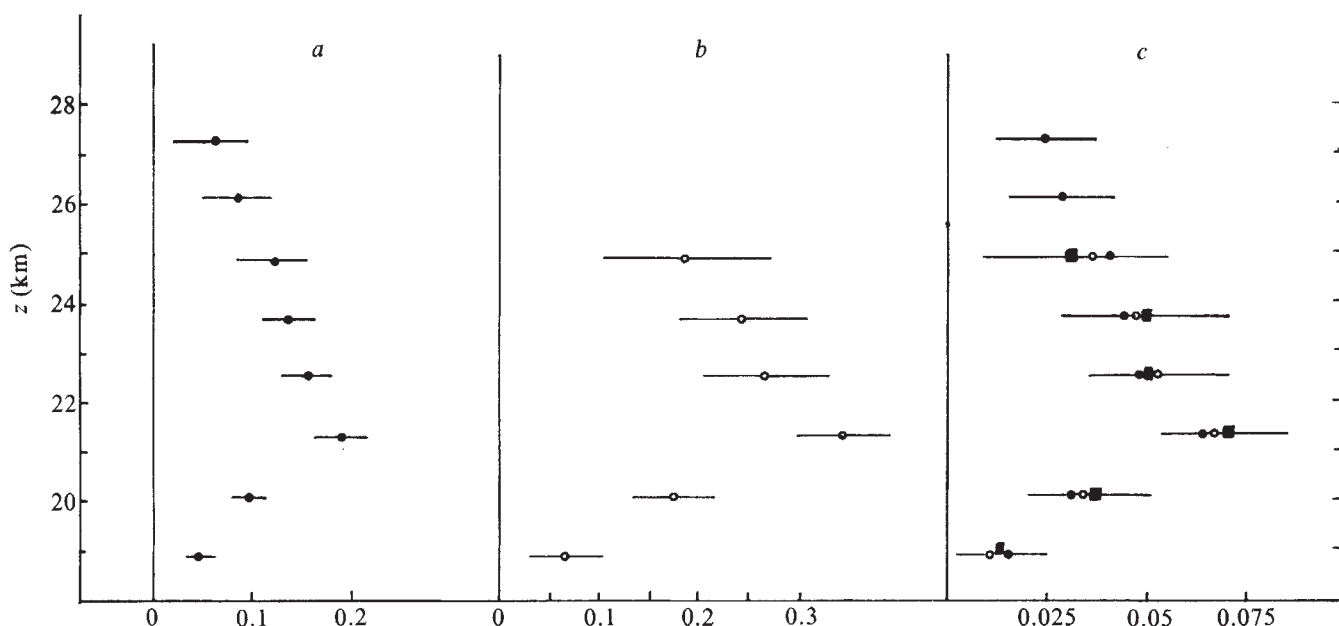


Fig. 2 a, Deviation of the differential optical thickness at $\lambda_1 = 308$ nm from a pure Rayleigh atmosphere ($\tau_{03} = 0$ corresponds to the situation of pure Rayleigh scattering). b, Deviation of the differential optical thickness at $\lambda_2 = 303.7$ nm from a pure Rayleigh atmosphere. ($\tau_{03} = 0$ corresponds to the situation of pure Rayleigh scattering). c, Comparison of $\tau_{03}(\lambda_1) - \tau_{03}(\lambda_2)$ as measured from equation (3), $\tau_{03}(\lambda_1)$ and $\tau_{03}(\lambda_2)$ normalised to the same absorption cross-section (10^{-19} cm 2). ($\tau/\sigma = 0$ corresponds to the situation of pure Rayleigh scattering.) (●, τ/σ (308.2 nm); ○, τ/σ (303.7 nm); ■, $(\tau(\lambda_1) - \tau(\lambda_2))/(\sigma(\lambda_1) - \sigma(\lambda_2))$).

both wavelengths and leads to the computed error bars, taking into account the statistical error of the received signals.

If we consider the ratio

$$\mathcal{R} = \frac{R(\lambda_1, z, z + \Delta z)}{R(\lambda_2, z, z + \Delta z)} = \exp -2[\tau_{O_3}(\lambda_1) - \tau_{O_3}(\lambda_2)] \quad (3)$$

we then obtain a direct determination of the ozone concentration without reference to any model. The backscattering by atmospheric particles has the same effect at two wavelengths and thus is eliminated. We have plotted on Fig. 2c the values of $(\tau_{O_3}(\lambda_1) - \tau_{O_3}(\lambda_2)) \times 10^{19} / (\sigma_{O_3}(\lambda_1) - \sigma_{O_3}(\lambda_2))$, together with the related error bars as computed from the received signal, assuming the following values for the absorption cross-section^{5,8}

$$\sigma_{O_3}(\lambda_1 = 308.2 \text{ nm}) = 3 \times 10^{-19} \text{ cm}^2$$

$$\sigma_{O_3}(\lambda_2 = 303.7 \text{ nm}) = 5.2 \times 10^{-19} \text{ cm}^2$$

These values are directly related to the ozone concentration and are compared with the ratios

$$\frac{\tau_{O_3}(\lambda_1) \times 10^{19}}{\sigma_{O_3}(\lambda_1)} \text{ and } \frac{\tau_{O_3}(\lambda_2) \times 10^{19}}{\sigma_{O_3}(\lambda_2)}$$

which lead to the same physical quantity. This shows good agreement between the two methods of data analysis confirming the former assumptions and the neutral density modelling.

The vertical ozone profiles obtained on these two nights are presented in Fig. 3 together as an example with the statistical error bar as derived for one night (20–21 July). The accuracy of the measurement is directly calculated from the backscattered signals using equation (3) except for the upper part of the profile ($z > 25 \text{ km}$) where the absorption at 303.7 nm is too strong to

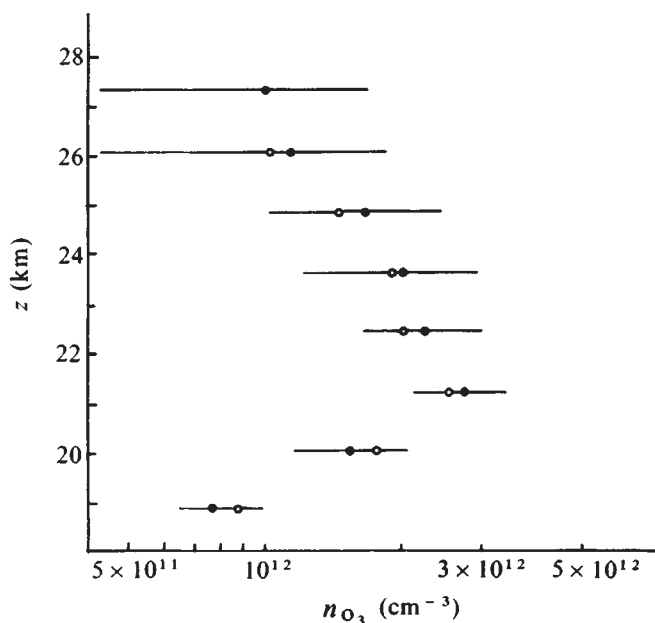


Fig. 3 Vertical ozone profiles (○, 19–20 July 1977; ●, 20–21 July 1977).

allow a precise determination of $\mathcal{R}$ using this channel. n_{O_3} is then obtained only from the 308 nm channel assuming a constant value for the atmospheric density scale height, as calibrated in the lower altitude range by comparison of $\tau_{O_3}(\lambda_1)$ and $\tau_{O_3}(\lambda_2)$. The accuracy of these preliminary measurements precludes any conclusion about a possible day to day variation of the ozone content. Its vertical repartition shows a maximum around 22 km with a concentration of $3 \times 10^{12} \text{ cm}^{-3}$ which is within the range of measurements by other methods⁹ and model computations¹⁰.

Further improvements of this DIAL technique to measure the ozone vertical profile will be done soon by extending the altitude range of the observations. Measurements of the backscattered signal in the analogical mode will allow the determination of n_{O_3} between 10–20 km whereas, in the lower troposphere, interference effects by other atmospheric minor constituents must be carefully considered. In a same way an increase of the output wavelength of the ultraviolet laser up to 320 nm will reduce the ozone absorption in the lower stratosphere and thus extend the altitude range of observations up to 40 km. Furthermore, the accuracy of the experiment could be improved by increasing the output energy of the ultraviolet laser up to a factor of 3 as already obtained in laboratory and its repetition rate (up to 2 Hz). This differential absorption lidar technique will then provide a very powerful tool for the determination of stratospheric ozone concentration from ground based stations.

G. MÉGIE
J. Y. ALLAIN
M. L. CHANIN
J. E. BLAMONT

Service d'Aéronomie du CNRS,
91370 Verrières-le-Buisson,
France

Received 10 August; accepted 3 October 1977.

1. Mégie, G., Bos, F., Blamont, J. E. & Chanin, M. L. *Planet. Space Sci.* (in the press).
2. Hirth, A., Vollrath, K. & Allain, J. Y. *Opt. Com.* **20**, 347 (1977).
3. Petit, R. E. W. *J. Atmos. Terr. Phys.* **37**, 669 (1975).
4. Gibson, A. J. & Thomas, L. *Nature* **256**, 561 (1975).
5. Ackerman, M. *Mesospheric Model and Related Experiments* (ed. Fiocco, G.) 149 (Reidel, Dordrecht, 1970).
6. Castleman, A. W., Jr *Physics and Chemistry of Upper Atmosphere* (ed. McCormac, B. M.) 143 (Reidel, Dordrecht, 1972).
7. U.S. Standard Atmosphere suppl. (1966).
8. Griggs, M. J. *Chem. Phys.* **49**, 857 (1968).
9. *The Natural Stratosphere* CIAP Monograph 1 (Washington, D. C. 1974).
10. Krueger, A. J. & Minzner, R. A. *J. geophys. Res.* **81**, 4477 (1976).

Photomineralisation rate of organic compounds adsorbed on particulate matter

ATTEMPTS to correlate the concentration of chemicals in the environment with their production figures have resulted in a large deficit, especially for organic chemicals such as DDT¹. It has been assumed that analytical errors accounted for this deficit. Another explanation, however, considers the reactions of compounds in abiotic conditions. As far as atmospheric chemistry is concerned, such an explanation needs a thorough knowledge of the kinetics and mechanisms of all possible reactions of a compound, including its products and also consideration of the meteorological conditions. Only then would it be possible to describe the time dependence of concentrations of organic compounds in the atmosphere with any accuracy. To consider all the data required is time-consuming and expensive, because of the rather complicated mechanisms of conversion and decomposition. We have therefore devised simple, standardised tests.

In previous experiments we demonstrated that even persistent chemicals were decomposed to CO_2 (and in the case of chlorinated compounds also to HCl and/or Cl_2) as solids in an oxygen stream^{2,3}. The results of this decomposition—the photomineralisation—are given in Table 1. Cyclodiene insecticides, such as aldrin and dieldrin, mineralise during irradiation with Pyrex-filtered ultraviolet light. Using short-wave ultraviolet light ($\lambda < 290 \text{ nm}$), the amount of mineralisation products is significantly increased. DDT, DDE and pentachlorophenol behave in the same way as aldrin and dieldrin, whereas results from irradiating pentachlorobenzene, hexachlorobenzene and certain chlorinated biphenyls (PCBs) lead us to assume that these compounds are only mineralised, when exposed to short-wave ultraviolet light. Photomineralisation is, therefore, a simple test evaluating the behaviour of chemicals in atmospheric conditions. The relative mineralisation rate which is used to compare one substance with others can be

Table 1 Mineralisation products after ultraviolet irradiation of certain chlorinated hydrocarbons as solids in an oxygen stream

Compound	Mineralisation products (mg)			
	Quartz (2 d)		Pyrex (6 d)	
	CO ₂	HCl	CO ₂	HCl
Aldrin				
Dieldrin	51-70	19-28	8-11	3-4
Photodieldrin				
Hexachlorobenzene				
Pentachlorobenzene				
2,4,5,2',4',5'-Hexachlorobiphenyl	46-53	19-26	†	†
2,5,2',5'-Tetrachlorobiphenyl				
Pentachlorophenol				
DDT				
DDE	*	*	10-15	2-8

About 80 mg of the respective compound was distributed on the inside wall of a 1-litre irradiation vessel and irradiated in an oxygen stream. The irradiation was carried out through quartz- and Pyrex-glass, using a Hg-high pressure lamp (HPK 125 W, Philips) as a light source.

*Not determined.

†Not detected.

easily determined by the mineralisation products formed, such as CO₂, HCl, Cl₂. But standardised and well controlled test conditions are necessary.

As a considerable fraction of the organic compounds in the atmosphere is adsorbed on particles, we use carrier materials in photomineralisation experiments and because of their readily standardisable properties, silica gel and aluminium oxide are good carriers. Using a suitable ratio of test substance to carrier material, a constant, approximately monomolecular dispersion on the surface can be assumed. The resulting substance-oxygen contact is more effective than that of solids and with the catalytic effects of the carrier material will increase the mineralisation rate. In addition the application of carrier materials was found to be particularly useful for comparing gaseous, liquid and solid exponents.

The photomineralisation of some chlorinated alkenes (1, 1-dichloroethylene, tetrachloroethylene, 2,3-dichloropropene, 1,1-dichloropropene and hexachlorobutadiene),

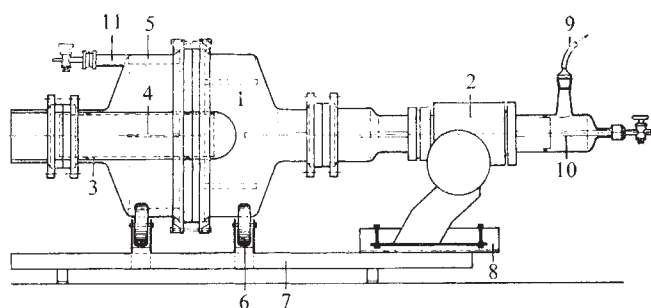


Fig. 1 Section of the irradiation apparatus for photomineralisation of adsorbed substances. The cylindrical reactor drum (1) is on one end linked by means of flanged joints, able to transmit power, with a drive motor (2) whose control is infinitely variable. On the other end a closed-ended irradiation tube (3) which is welded with a glass coupling and attached to the reactor drum by a coupling, projects into the interior of the reactor. The dimensions of the irradiation tube are chosen so that different sources of light (4) and filters can be used. On the inner wall of the reactor drum blades (5), parallel to the longitudinal axis, are inserted, which occupy only a fraction of the diameter and are adjusted into staggered positions starting from the front end. The weight of the reactor drum is carried by four rollers (6), which are connected to a support (7) by a holding device. On the support the flexible suspension for the drive motor is arranged. The reactor drum can be charged either through tube (10) at the stationary part of the apparatus or through tube (11) by applying a vacuum to tube (9). By rotation of the reactor drum along its longitudinal axis the reactants (carrier material with adsorbed phase and gas phase) are completely tumbled and mixed so that a homogeneous irradiation is guaranteed.

freons (trichlorofluoromethane and dichlorodifluoromethane) and aromatic hydrocarbons (benzene, toluene, and *ortho*-, *meta*- and *para*-xylenes) in the adsorbed state was carried out in a special irradiation apparatus (Fig. 1)⁴. About 1.3 g of the test substance was applied to 800 g of silica gel using a stream of nitrogen gas. The amount of substance actually adsorbed was determined by gas chromatography and by elementary analysis. Silica gel (350 g) containing the test substance were introduced into the irradiation drum which was then evacuated twice to 100 mm Hg and flooded with pure oxygen (at a drum capacity of about 25 l, a large excess of oxygen is ensured in respect to a total mineralisation of the initial amount adsorbed). The ultraviolet irradiation of the adsorbed substances was carried out with a Hg-high pressure lamp (HPK 125 W, Philips) using filters of quartz or Pyrex glass for 3 or 6 d respectively. After the irradiation the mineralisation products, such as CO₂, Cl₂ were driven off from the reactor drum by a carrier gas (N₂) and trapped by passing the gas through two wash bottles connected in series, each containing 50 ml of 0.1 M NaOH solution. CO₂ was quantitatively determined by acidometric titration of CO₃²⁻ and Cl₂ by iodometric titration of OCl⁻. The quantitative determination of mineralisation products which could not be eluted from the silica gel (such as HCl) was done by mixing silica gel with a NaOH solution (0.1 M) and subsequent potentiometric titration of Cl⁻ with AgNO₃. Table 2 and Fig. 2 show the results of the photomineralisation of the chlorinated alkenes, freons and aromatic hydrocarbons. In

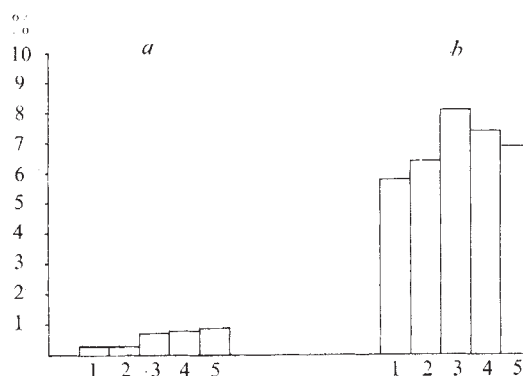


Fig. 2 CO₂-formation of aromatic hydrocarbons adsorbed on silica gel by ultraviolet-irradiation. Percentages are given by moles CO₂ formed divided by moles CO₂ expected from a total mineralisation of the initial amount adsorbed. *a*, Pyrex, 6 d; *b*, quartz, 4 d. 1, Benzene; 2, toluene; 3, *o*-xylene; 4, *m*-xylene; 5, *p*-xylene.

all cases CO₂ was found; in the case of chlorinated compounds HCl and/or Cl₂ were also found.

The photomineralisation of chlorinated alkenes is the most effective as shown in Table 2. The quantity of the formed mineralisation products is not noticeably influenced by using quartz- or Pyrex-filtered ultraviolet light. The increased amount of products in the Pyrex experiments are due to the extended irradiation times. This indicates a possible photomineralisation with ultraviolet light predominating in the troposphere ($\lambda > 290$ nm). Even the freons which are well known as highly persistent compounds are degraded to CO₂ and HCl by using Pyrex-filtered light. But here the degradation with short-wave ultraviolet light is slightly favoured.

In contrast to the chlorinated alkenes and freons, the aromatic compounds in the solid state and adsorbed on silica gel show a higher persistence with respect to a photomineralisation. While irradiating through Pyrex glass only low mineralisation rates were found. Using quartz-filtered

Table 2 Photomineralisation of chlorinated alkenes and freons

	Quartz (3 d)		Pyrex (6 d)	
$\text{CCl}_2 = \text{C} - \text{C} = \text{CCl}_2$	*	§	†	§
$\begin{array}{c} \text{Cl} \quad \text{Cl} \\ \quad \\ \text{CH}_3 - \text{CH} = \text{CCl}_2 \\ \quad \\ \text{H}_2\text{C} - \text{C} = \text{CH}_2 \end{array}$	†	‡	†	§
$\begin{array}{c} \text{Cl} \quad \text{Cl} \\ \quad \\ \text{Cl}_2\text{C} = \text{CCl}_2 \end{array}$	*	‡	†	§
$\text{H}_2\text{C} = \text{CCl}_2$	*	‡	†	§
CCl_3F	*	§	*	‡
CCl_2F_2	*	‡	*	‡

Percentages are given by the amounts of CO_2/Cl actually formed divided by the amounts of CO_2/Cl expected from a total mineralisation of the initial quantity adsorbed.

*10–50% HCl - and/or Cl_2 -formation.

†50–90% HCl - and/or Cl_2 -formation.

‡10–50% CO_2 -formation.

§50–90% CO_2 -formation.

light, however, a significant increase could be determined (Table 1 and Fig. 2).

An even greater distinction within one class of compounds is possible: by comparing benzene with toluene and the xylenes (Fig. 2) a tendency to higher CO_2 -formation was found. Moreover, the three xylenes showed differences due to the steric position of their methyl groups. These results are supported by the determination of the remaining amounts of aromatic hydrocarbons after irradiation.

From these examples a characteristic behaviour of certain compounds as well as classes of substances becomes evident, which might help in evaluating organic chemicals with regard to their degradation in environmental conditions.

S. GÄB
J. SCHMITZER
H. W. THAMM
H. PARLAR
F. KORTE

Institut für Ökologische Chemie der Gesellschaft für Strahlen- und Umweltforschung mbH München, D-8051 Attaching, FRG
and

Institut für Chemie der Technischen Universität München, D-8050 Freising-Weihenstephan, FRG

Received 5 August; accepted 5 October, 1977.

1. Woodwell, G. M., Craig, P. P. & Johnson, H. A. *Science* **174**, 1102–1106 (1971).
2. Gäb, S., Parlar, H., Nitz, S., Hustert, K. & Korte, F. *Chemosphere* **3**, 183–186 (1974).
3. Gäb, S., Nitz, S., Parlar, H. & Korte, F. *Chemosphere* **4**, 251–256 (1975).
4. *Deutsche Patentanmeldung* P 27 04 942.9.

Formation of Fe–Ni–Si planetary cores

THE planets may have accreted inhomogeneously materials accumulating on the proto-planets in the sequence in which they condense at high temperatures (1,000–2,000 K) out of the solar nebula¹. The first solids to condense from the Earth, however, are the Ca and Al oxides, silicates and titanates^{2,3} and not Fe or Fe–Ni alloy, which is supposed to form the Earth's core. Anderson and Hanks⁴ consider that such a nucleus of oxides enriched in radioactive elements could provide a mechanism for heating and melting the surrounding solid Fe layer which accretes later. Ringwood⁵ considers this model of core formation unsatisfactory since it may produce a layer of oxide rich material around the core for which we have no geophysical evidence. For Fe, or an alloy of Fe with other

elements, to form as the first equilibrium solid, the pressure in the solar nebula at 1,800 K must be several atmospheres^{3,6}. But, the conditions for the formation of a liquid Fe or a liquid Fe-rich alloy have not been determined. If a liquid alloy formed as the first condensate, it would solve many problems in the current models of formation of planetary cores⁵. The pressure required to form a liquid alloy is likely to be an order of magnitude lower than the pressure for the formation of solid Fe because metal liquid solutions of Fe, Ni and Si have a lower free energy of solution than that of an ideal solution. In particular the Fe-liquid and Si-liquid, and Ni-liquid and Si-liquid form binary solutions with a low free energy which stabilises the liquid solutions at pressures lower than that of the solid. According to Podolak and Cameron⁶ the models of Jupiter have an enhanced O–H ratio of ~25–30. They consider that pressure required to form chondrules will be lowered several orders of magnitude if the O/H ratio is increased. We report here an investigation into the equilibrium condensation of the solar nebular gas at 1,900 K at varying total pressures and varying abundance of hydrogen (A_H).

The system consisted of 10 elements H, O, Al, Mg, Fe, Ni, C, Si, S, Ca in solar abundance⁷ and over 70 gaseous and condensed compounds for which thermochemical data are available in sources listed by previous workers (such as Grossman³). The condensed compounds included forsterite and a number of liquid solutions of Fe, Ni and Si. These three elements were selected because they are liquids at 1,900 K. The vapour pressures of all other elements are low at 1,900 K and therefore none dissolves in the melt in any significant concentration. None of the oxides, titanates and other silicates condense at 1,900 K.

The free energies of the formation of liquid solutions were calculated using the expression

$$G = RT \sum X_i \ln X_i \gamma_i$$

Where G is the free energy of formation of the solution, X_i the mole fraction, γ_i the activity coefficient, R the gas constant and T the absolute temperature. For a ternary solution, γ_i is given by⁸:

$$\ln \gamma_i = X^2 [E_{ij} + 2 X_i (E_{ji} - E_{ij})] + X_k^2 [E_{ik} + 2 X_i (E_{ki} - E_{ik})] + X_{jk} [0.5 (E_{ji} + E_{ij} + E_{ki} + E_{ik} - E_{jk}) + X_i (E_{ji} - E_{ij} + E_{ki} - E_{ik}) + (X_j - X_k) (E_{jk} - E_{kj}) - (1 - 2X_i) C]$$

Where the E values are the binary interaction coefficients and C is the ternary constant. Here we neglect the ternary constant. The data on binary coefficients are $E_{\text{Fe-Si}} = -5.77$, $E_{\text{Si-Fe}} = -7.41$, $E_{\text{Ni-Si}} = -6.75$, $E_{\text{Si-Ni}} = -19.43$, $E_{\text{Fe-Ni}} = -0.67$ and $E_{\text{Ni-Fe}} = -0.67$. These coefficients have been calculated from the thermochemical data given by Fruehan⁹ (Fe–Si), by Kubachewski *et al.*¹⁰ (Ni–Si) and by Elliot *et al.*¹¹ (Fe–Ni).

The results of the calculations are shown in Fig. 1 and Table 1. Following the practice of Herndon and Suess¹², in Fig. 1, we used the hydrogen depletion factor (here defined as $D_H = \log A_H$ in solar gas/ A_H chosen to vary the ratio; no other element is depleted). Under most conditions of pressure, and varying A_H the liquid that forms does not differ appreciably in composition from $\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$. At normal hydrogen abundance, the liquid does not form until the pressure exceeds 1 atm. With decreasing A_H ($D_H = 1$) the minimum pressure at which liquid forms is between 0.1 and 0.2 atm. The amount of liquid increases rapidly with increasing pressure and by 1 atm most of the Fe and Ni have condensed along with forsterite (Mg_2SiO_4). A significant amount of Fe and Ni (~60%) condense as liquid solution without any admixture with forsterite between 0.1 and 0.3 atm. Therefore if such conditions of pressure and hydrogen depletion could exist in the nebula over a length of time an alloy core with an outer layer of alloy and forsterite could form.

With a further decrease in A_H ($D_H = 2$) the minimum pressure for liquid formation lies between 0.01 and 0.05 atm. This liquid, however, does not contain any Si. The concentration of

Table 1 Condensed species forming at 1,900 K

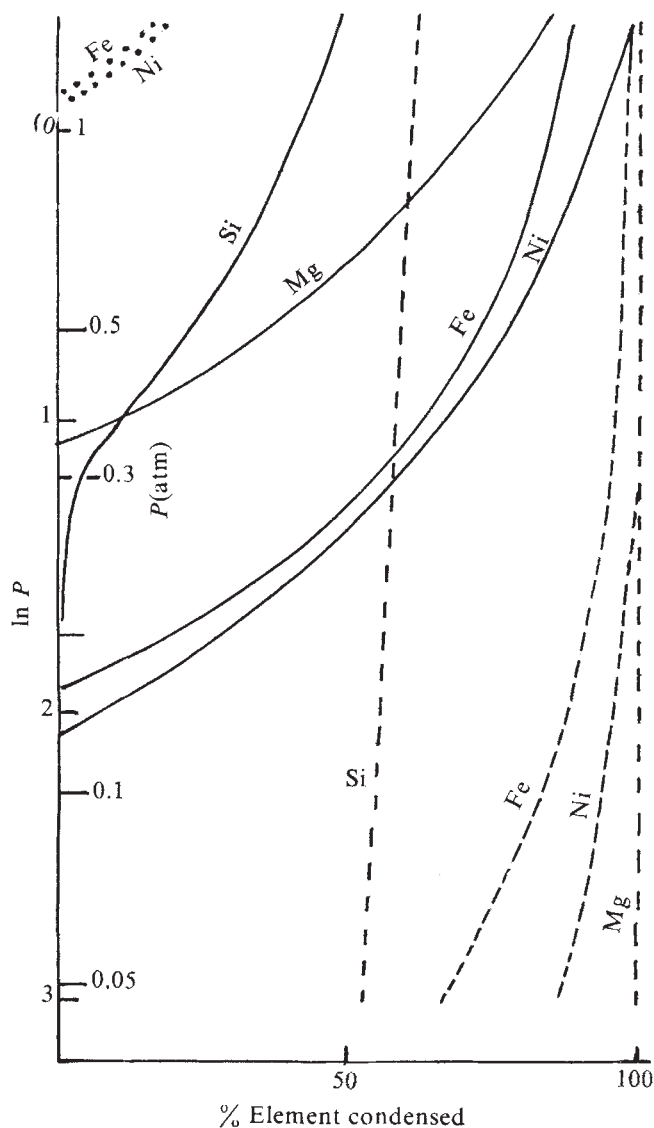
Pressure (atm)	$A_H = 2.6 \times 10^{10}$		$A_H = 2.6 \times 10^9$		$A_H = 2.6 \times 10^8$	
	Liquid	Solid	Liquid	Solid	Liquid	Solid
1.5	$\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$	None	$\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$	Mg_2SiO_4	$\text{Fe}_{0.906}\text{Ni}_{0.047}\text{Si}_{0.047}$	Mg_2SiO_4
1.0	None	None	$\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$	Mg_2SiO_4	$\text{Fe}_{0.905}\text{Ni}_{0.046}\text{Si}_{0.047}$	Mg_2SiO_4
0.5	None	None	$\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$	Mg_2SiO_4	$\text{Fe}_{0.903}\text{Ni}_{0.045}\text{Si}_{0.048}$	Mg_2SiO_4
0.4	None	None	$\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$	Mg_2SiO_4	$\text{Fe}_{0.902}\text{Ni}_{0.044}\text{Si}_{0.049}$	Mg_2SiO_4
0.3	None	None	$\text{Fe}_{0.9}\text{Ni}_{0.05}\text{Si}_{0.05}$	None	$\text{Fe}_{0.948}\text{Ni}_{0.052}\text{Si}_{0.03}$	Mg_2SiO_4
0.2	None	None	$\text{Fe}_{0.9}\text{Ni}_{0.06}\text{Si}_{0.04}$	None	$\text{Fe}_{0.948}\text{Ni}_{0.052}$	Mg_2SiO_4
0.1	None	None	None	None	$\text{Fe}_{0.945}\text{Ni}_{0.055}$	Mg_2SiO_4
0.05	None	None	None	None	$\text{Fe}_{0.938}\text{Ni}_{0.062}$	Mg_2SiO_4
0.01	None	None	None	None	None	Mg_2SiO_4

A_H , hydrogen abundance. The compositions of the liquid solutions are approximate.

Si in liquid increases with pressure. At such a hydrogen abundance, Mg is nearly 100% condensed as forsterite at all pressure considered in this study.

Our conclusions are: (1) A core of liquid alloy could form from a gas of solar composition only at relatively high pressures of the order of 1 to 2 atm. (2) This pressure could be lowered by

Fig. 1 Percentage of elements condensed based on the results of phase equilibrium calculations. Fe, Ni and Si condense as liquid solution, Mg and Si as Mg_2SiO_4 . D_H values as 0 (dotted line), 1 (solid line) and 2 (dashed line) represent hydrogen abundances of 2.6×10^{10} , 2.6×10^9 and 2.6×10^8 respectively.



about a factor of 10 if hydrogen is depleted by a similar factor. (3) A greater hydrogen depletion ($D_H = 1$) results in stabilising forsterite down to 0.01 atm (and lower). Significant amount of liquid ($\text{Fe}_{0.94}\text{Ni}_{0.06}$) forms only at 0.5 atm which could be regarded as the lowest pressure for the formation of the liquid. (4) Silicon may be the light element component of the alloy only if pressures are higher than 0.2–0.3 atm. In the pressure range considered (0.2 to 1.5 atm) the concentration of silicon in the liquid solution varies between 3 and 5%. If the O/H conditions can be deduced from other mineral–chemical information, an appreciable concentration of Si in the metal phase could indicate formation at high nebular pressure.

We thank Drs A. E. Ringwood and R. Brett for helpful suggestions. Part of this work was done while S.K.S. was visiting Dr M. J. Drake at the Lunar and Planetary Laboratory, University of Arizona, Tucson. Phase equilibria was computed using a program from Dr G. Eriksson, University of Umeå, Sweden.

SURENDRA K. SAXENA
A. BENIMOFF

Department of Geology,
City University of New York,
Brooklyn College,
Brooklyn, New York 11210

Received 6 May; accepted 21 September 1977.

1. Turekian, K. K. & Clark, S. P., Jr *Earth planet. Sci. Lett.* **6**, 346 (1969).
2. Larimer, J. W. & Anders, E. *Geochim. cosmochim. Acta* **31**, 1239 (1967).
3. Grossman, L. *Geochim. cosmochim. Acta* **36**, 597 (1972).
4. Anderson, D. L. & Hanks, T. C. *Nature* **237**, 387 (1972).
5. Ringwood, A. E. *Composition and petrology of the earth's mantle* (McGraw-Hill, New York, 1975).
6. Podolak, M. & Cameron, A. G. W. *Icarus* **23**, 326 (1974).
7. Cameron, A. G. W. in *Origin and Distribution of the Elements* (ed. Ahrens, L. H.) 125 (Pergamon, Oxford, 1968).
8. King, M. B. *Phase Equilibrium in Mixtures* (Pergamon, Oxford, 1969).
9. Fruehan, R. J. *Met. Trans.* **1**, 865 (1970).
10. Kubachewski, O., Evans, E. L. I. & Alcock, C. B. *Metallurgical Thermochemistry* (Pergamon, Oxford, 1967).
11. Elliott, J. F., Gleiser, M. & Ramakrishna, V. *Thermochemistry for Steel-making* 2 (Addison-Wesley, Reading, Massachusetts, 1963).
12. Herndon, J. M. & Suess, H. E. *Geochim. cosmochim. Acta* **41**, 233 (1977).

Angle of subduction

PLATE tectonics is well established as an empirical description for modification of the Earth's outermost solid layer, although the dynamics of the plates are poorly understood. One view is that the plates are surface manifestations of the deep mantle convection which must inevitably occur if there are deep-seated energy sources¹. Mantle convection, however, may not be an efficient driver of plate motions², and the negative buoyancy of subducting slabs may be more important^{2, 3}. We consider here a simple fluid dynamical model which illustrates some of the general principles of subduction dynamics, and suggest an explanation for the observed subduction angle (the angle between the subducting slab and the Earth's surface). The model depends on the concept of a 'critical' negative buoyancy that plate material must attain before it can undergo steady-state subduction.

Observations indicate that the subduction angle θ_s is seldom near 0° or 90° but is typically an intermediate value, say 50° (refs 4, 5). This does not seem to be explainable by trench migration (or the relative motion of plates) alone⁵. Jischke⁶ explained the angle by proposing that the subducting slab adheres to the overriding continent, so that its angle of descent is dictated by the continental geometry, but this is only satisfactory for near surface behaviour. A detailed force-balance or energy-balance model might correctly predict θ_s , but only by a fortuitous choice of poorly known parameters such as the average mantle viscosity μ . In the energy minimisation calculation by Tullis⁵, for example, there is only a narrow range of μ for which the predicted θ_s is not near 0° or 90° . We propose, instead, that θ_s is essentially independent of the poorly known mantle viscosity and plate rheology, and is determined by the requirement that the torques on the slab are balanced for the lowest possible negative buoyancy. The highly simplified model that we consider predicts $\theta_s \sim 63^\circ$ for a non-migrating trench. Slab rigidity is fundamental to our model, but since the forces acting on the slab are a function of depth, deformation and even fracture of the slab may eventually occur and play a role in mantle seismic mechanisms⁷. Torque balance, however, with an appropriately chosen slab length, is a suitable first approximation for estimating θ_s .

For simplicity, we assume that the mantle has a uniform viscosity and is convectively neutral, so that the only motion within it is that caused by the plates. We consider a two-dimensional model in which the trench and upper (non-subducting) plate are at rest (see Fig. 1). The subducting slab is assumed to be rigid and moving at a steady speed U . For the purpose of calculating the induced mantle flow only, it is also assumed to be infinitely long and infinitesimally thin. This is

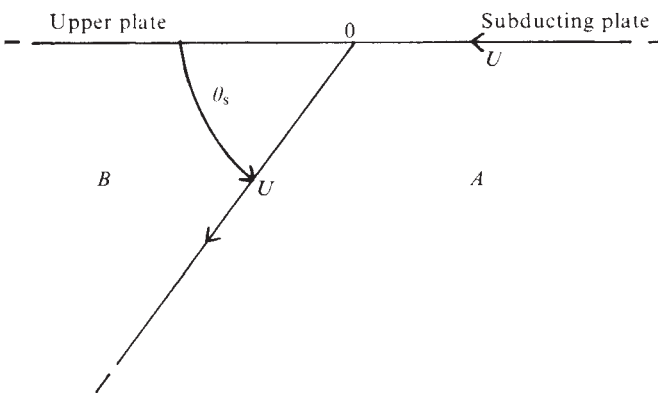


Fig. 1 Two-dimensional model for a subduction zone (see ref. 7). Material is transported from the ridge (to the right of the diagram) and undergoes subduction at 0. The upper plate is at rest.

the model considered by McKenzie⁸, whose results we use here. He gives the hydrodynamic contribution to the pressure in regions A and B of Fig. 1 are $P_A(\theta)$, and $P_B(\theta)$ respectively, where

$$P_A(\theta) = \frac{2\mu U [\sin \theta - \sin(\theta - \theta_s)]}{r[(\pi - \theta_s) + \sin \theta_s]}$$

$$P_B(\theta) = \frac{-2\mu U [\sin \theta \sin \theta_s + \theta_s \sin(\theta_s - \theta)]}{r[\theta_s^2 - \sin^2 \theta_s]} \quad (1)$$

θ is the polar coordinate measured anticlockwise from the upper plate, and r is the radial coordinate measured from 0. The lubrication resulting from shear heating⁸ and the diapiric upwellings which are responsible for behind-arc spreading⁹ both modify these pressure estimates at small r , but this does not greatly affect the torque balance that we consider.

Since $P_A(\theta_s) > 0 > P_B(\theta_s)$ for all θ_s , the hydrodynamic pressure difference across the slab always acts to reduce θ_s . The torque about 0 due to this suction is T_H per unit trench length, where

$$T_H = \int_0^L [P_A(\theta_s) - P_B(\theta_s)] r dr$$

$$= 2\mu UL \left[\frac{\sin \theta_s}{(\pi - \theta_s) + \sin \theta_s} + \frac{\sin^2 \theta_s}{\theta_s^2 - \sin^2 \theta_s} \right] \quad (2)$$

and L is the length of the subducting slab (assumed independent of θ_s). We assume that T_H is balanced by the gravitational torque T_G caused by the negative buoyancy of the slab. If $\Delta\rho(r)$ is the density contrast between slab and mantle, and h is the slab thickness, then

$$T_G = \int_0^L \Delta\rho(r) g h r \cos \theta_s dr$$

$$= \frac{1}{2} b L^2 \cos \theta_s$$

$$b = g h \int_0^1 \Delta\rho(x) dx \quad (3)$$

where $x = r/L^2$, g is the gravitational acceleration, and b is the mean negative buoyancy per unit slab area. The inclusion of h in our definition of b emphasises that subduction is aided either by increasing the density contrast or by thickening the plate. We neglect the torque required to maintain the bend *a priori*, it should be noted that whereas the work done at the trench may be large, it does not follow that the corresponding torque on the subducting slab is necessarily large.

The graphical solution of $T_H = T_G$ is shown in Fig. 2. Since $T_H > 0$ for $0 \leq \theta_s \leq \pi$ it follows that, for a fixed U , there is no solution for b sufficiently small. A solution becomes possible at a critical negative buoyancy b_c and a critical subduction angle θ_c when the curves for T_H and T_G touch. The critical solution is

$$\theta_c \simeq 63^\circ, b_c \simeq 20\mu U/L \quad (4)$$

For $b > b_c$, two solutions are possible. The larger solution for θ_s is stable since any small perturbation from it clearly leads to a restoring torque. The smaller solution for θ_s is unstable, since any small deviation from it is amplified. The simple model predicts that $\theta_s < 63^\circ$ is never stable. Subduction at these angles may be possible because of the finite torque required to bend the slab or the hydrodynamic pressure on the leading edge of the slab².

This calculation provides no information on the onset of subduction since it concerns a steady state for which a well-developed subduction slab already exists. A model by Davies¹⁰ may partly explain the initial subduction angle. We also do not consider the force balance along the direction of subduction which determines U . This balance is affected not only by the negative buoyancy and hydrodynamic effects², but also by viscous drag on the subducting plate and ridge pushing (gravitational sliding)¹¹. It is likely that U is not strongly dependent on θ_s , but must itself exceed a critical value that is determined by the requirement that shear heating produce adequate lubrication below the plate between ridge and trench^{12, 13}.

The following speculative but plausible picture emerges: after subduction is initiated, the trench migrates until a steady-state is reached where newly subducting material has the critical negative buoyancy b_c and subducts at an angle θ_c .

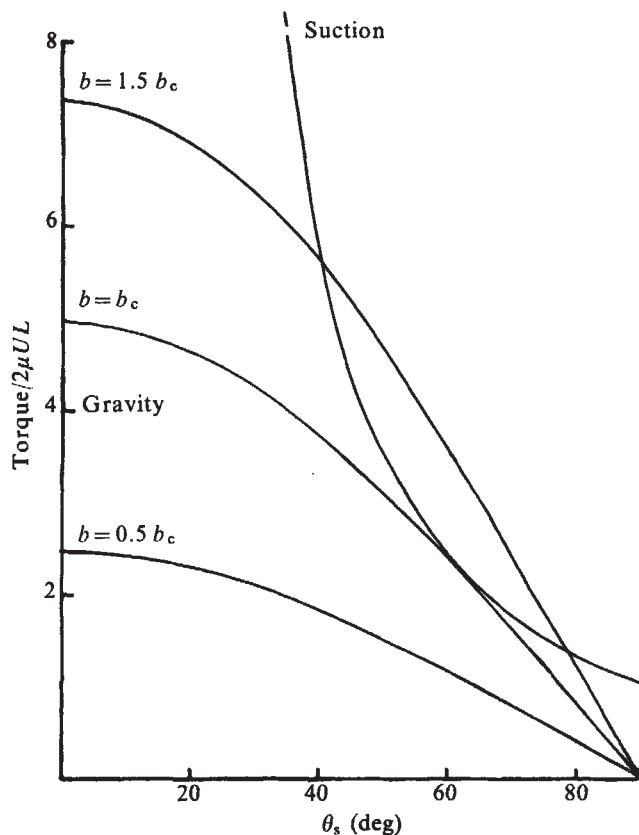


Fig. 2 The torques (in units of $2\mu UL$) as a function of subduction angle. The three cosine curves labelled by values of b are for T_G . The remaining curve is for T_H . There is no solution for $T_H = T_G$ if $b < b_c$, and a solution becomes possible at $b = b_c$ when the curves just touch.

The negative buoyancy results from the thickening and cooling of the plate as the material proceeds from ridge to trench. The observed scatter of actual values for θ_s results from the inevitable interaction between several ridge and trench systems, and the inability to achieve steady-state. The following plausible values satisfy equation (4) for b_c : $\Delta\rho \sim 0.1 \text{ g cm}^{-3}$, $h \sim 100 \text{ km}$, $U \sim 10 \text{ cm yr}^{-1}$, $L \sim 1,000 \text{ km}$, $\mu \sim 2 \times 10^{22} \text{ P}$. The size of the plate depends on the time taken for plate material to achieve critical negative buoyancy and, therefore, depends on the geothermal heat flux. This dependence may help to explain why plate tectonics was apparently so different in the Precambrian¹⁴.

The model we present here is oversimplified and incomplete, but it illustrates how the general physical principles of subduction dynamics might be identified. The choice of model is best made by comparing different tectonic systems, and data on the plate tectonics for the Earth at a different epoch or for other planets is essential for future progress.

After this work had been completed, the very closely related calculations by Tovish *et al.*¹⁵ were brought to our attention.

We thank A. L. Hales and A. E. Ringwood for useful comments.

D. J. STEVENSON
J. S. TURNER

Research School of Earth Sciences,
Australian National University, Canberra, Australia 2600

Received 11 July; accepted 20 September 1977.

1. Turcotte, D. L. & Oxburgh, E. R. *An. Rev. Fluid Mech.* 4, 33–68 (1972).
2. Richter, F. M. *Tectonophysics* 38, 61–88 (1977).
3. Forsyth, D. & Uyeda, S. *Geophys. J. R. astr. Soc.* 43, 163–200 (1975).
4. Luyendyk, B. P. *Geol. Soc. Am. Bull.* 81, 3411–3416 (1970).
5. Tullis, T. E. Preprint (1977).
6. Jischke, M. C. *J. Geophys. Res.* 80, 4809–4813 (1975).
7. Isaacs, B. & Molnar, P. *Nature* 228, 1121–1124 (1969).
8. McKenzie, D. P. *Geophys. J. R. astr. Soc.* 18, 1–32 (1969).
9. Oxburgh, E. R. & Parmentier, E. M. *J. Geol. Soc., Lond.* 133, 343–355 (1977).

10. Davies, G. F. *Geophys. J. R. astr. Soc.* 49, 557–563 (1977).
11. Hales, A. L. *Earth planet. Sci. Lett.* 6, 31–34 (1969).
12. Froidevaux, C. & Schubert, G. *J. Geophys. Res.* 80, 2553–2564 (1975).
13. Melosh, H. J. *Tectonophysics* 35, 363–390 (1976).
14. McElhinny, M. W. & McWilliams, M. O. *Tectonophysics* 40, 139–159 (1977).
15. Tovish, A., Luyendyk, B. & Schubert, G. *Trans. Am. Geophys. Un.* EOS 57, 1002 (1976).

Age and tectonic implications of the Baie d'Audierne basic-ultrabasic complex

MAJOR deformational and metamorphic events in the south-armoricain metamorphic belt of south Brittany have been considered to be either Cadomian and Variscan¹ or Ordo-Silurian (Ligerian) and Variscan² in age. Geological interest in southern Brittany is, at least in part, focused upon the blueschists³ and the basic-ultrabasic complexes of the area. The largest of these basic-ultrabasic complexes is that of the Baie d'Audierne group (Fig. 1). Although lacking suitable material for direct radiometric or stratigraphical dating, several indirect attempts have been made to date this group, using field data. These are briefly outlined here. It is proposed that the deformational and metamorphic history of this complex is entirely Variscan in age. The implications of this conclusion for the recently proposed plate-tectonic model of Cogné² are discussed.

The south-armoricain metamorphic belt comprises a suite of non-fossiliferous, pre-Stephanian sediments, commonly ascribed to the late Proterozoic Brioverian⁴, which have been folded, metamorphosed and intruded by basic and ultrabasic rocks, granites and granite gneisses^{5,6}. In the Baie d'Audierne region (Fig. 1), two divisions within the metamorphic rocks, an upper and a lower suite separated by an unconformity have been recognised^{7–9}. It is thought that the lower suite was folded and metamorphosed in the amphibolite facies (kyanite grade) during the late Proterozoic Cadomian orogeny⁸ before the deposition of the upper suite. The upper suite commenced with the deposition of a basal sedimentary series, the Languidou gneiss. Ordovician to Carboniferous sediments lie offshore of the Baie d'Audierne; their relationships to the upper suite is, however, uncertain⁹. In the Baie d'Audierne, the upper suite is cut by the Pors Poulhan granite gneiss (Fig. 1) which has yielded a whole rock Rb/Sr isochron at $334 \pm 8 \text{ Myr}$ ($\lambda^{87}\text{Rb} = 1.47 \times 10^{-11} \text{ yr}^{-1}$). This radiometric date has been interpreted as the magmatic cooling age of the granite. Subsequently, the whole sequence, upper and lower suites and the intruded Pors Poulhan granite, was folded by upright folds on east-west axes and metamorphosed in the albite-amphibolite facies during the Variscan orogeny⁸. Late Variscan granites intrude all the metamorphic rocks and are probably associated with other south-armoricain granites emplaced into rocks of Upper and Lower Palaeozoic age¹¹.

During recently completed fieldwork in the area where Cogné and Peucat defined the upper and lower Brioverian, in the Baie d'Audierne region (Fig. 1), I was unable to recognise any unconformity. The upper and lower suites throughout the area show a common structural and metamorphic history. In addition, the Languidou gneiss, interpreted by Cogné⁷ as a metaconglomerate and used by Peucat⁸ to support the postulated unconformity, is thought to be an intrusive granite containing partially assimilated xenoliths of country rock. The supposed 'pebbles' are deformed feldspar porphyroblasts and quartz-feldspar segregations. The Languidou gneiss shows the same pre-tectonic relationship to the surrounding metasediments as the Pors Poulhan gneiss. This pre-tectonic relationship also applies to an association of gabbro and peridotite (now amphibolite, garnet-pyroxenite and serpentinite, ascribed by Peucat to his lower suite) which intrude the metasediments. Following the intrusion of these ultrabasic to granitic rocks, the entire complex (the Baie d'Audierne group⁶) was isoclinally folded and metamorphosed in the amphibolite facies, up to sillimanite grade, with the development of an axial planar schistosity. This early schistosity was deformed by rare isoclinal folds and

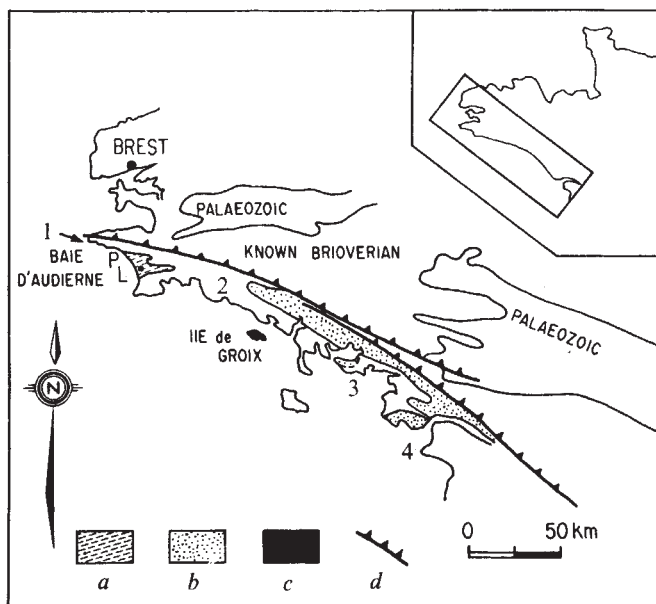


Fig. 1 Present-day distribution of the Baie d'Audierne group (a), south Brittany migmatites, (b) the Ile de Groix blueschists (c) and the spatial relationship to the Brioverian and Palaeozoic sediments north of the South Armorican Shear Zone (d); the blueschists may originally have been more widely separated from the mainland migmatites as they are bounded north and south by faults, possibly thrusts²⁴. L, Languidou; P, Pors-Poulhan. 1, Cap Sizun gneisses; 2, Port Manech gneiss; 3, Morbihan; 4, Basse Loire (After Cogné¹).

these in turn are refolded by upright folds on east-west axes, with a variably developed associated crenulation cleavage. The complex then underwent a mimetic annealing recrystallisation which did not produce any major retrogression of the earlier mineral assemblages. A subsequent retrogression in the greenschist facies is associated with the emplacement of the latest Variscan granites. A fuller account of the Baie d'Audierne area is given elsewhere⁶.

Similarly, the structural discordance representing the lateral extension of the unconformity within the metasediments of the metamorphic belt south of the South Armorican Shear Zone has been observed neither in Morbihan¹² nor Basse Loire¹³ (Fig. 1). From the above model, all the structural and metamorphic events affecting the rocks of the Baie d'Audierne group occurred between the emplacement of the Pors Poulhan gneiss and that of the late Variscan granites. If we accept the 334 ± 8 Myr age given by the Pors Poulhan gneiss as an emplacement age, then the folding and metamorphism in the Baie d'Audierne area must have occurred entirely within the Upper Palaeozoic and forms part of the Variscan orogeny.

To the east, south east of the Baie d'Audierne, a suite of migmatites and anatectic granites outcrops (Fig. 1), within which the anatectic granite represents the culmination of migmatization¹. To the south of the migmatites, metamorphic rocks of the blueschist (glaucofane-lawsonite) facies outcrop on the Ile de Groix^{3,14,15}. The anatectic granites have yielded a whole rock Rb/Sr isochron at 363 ± 12 ($\lambda^{87}\text{Rb} = 1.47 \times 10^{-11} \text{ yr}^{-1}$) (ref. 16). A premigmatization orthogneiss (Roguedas) has yielded a Rb/Sr whole rock isochron age at 460 Myr ($\lambda^{87}\text{Rb} = 1.42 \times 10^{-11} \text{ yr}^{-1}$) (ref. 17), so fixing the migmatization between 460 Myr and 363 Myr. The diapiric rise of the anatectic granites is syn-second phase deformation in the surrounding migmatites¹⁸. This, in conjunction with the radiometric data, suggests the presence of a polyphase history of pre-Variscan², high temperature metamorphism and deformation to the east whose early events and climax are older than the proposed post-334 Myr orogenic activity in the Baie d'Audierne to the west.

The radiometric ages for the migmatite suite are broadly similar to those for the time range of blueschist metamorphism

on the Ile de Groix (420–370 Myr) (ref. 17) as inferred from Rb/Sr mineral and whole rock data ($\lambda^{87}\text{Rb} = 1.42 \times 10^{-11} \text{ yr}^{-1}$). A K/Ar 335 Myr mineral age has been suggested for this blueschist metamorphism¹⁹, which may, however, represent a later thermal event. In view of the spatial association of the above mentioned radiometric ages, rock types and metamorphic facies, it is possible to place the rocks of the Baie d'Audierne within the context of recently proposed tectonic models for south Brittany^{2,20}. Such a tectonic context is necessarily confined to the south of the South Armorican Shear Zone due to the uncertainty concerning the movement history on this major tectonic line (P. Jegouzo, personal communication) and, therefore, the questionable validity of direct correlation between the south armorian metamorphic belt and central or north Brittany, the British Isles and so on.

The dominant strike trend in the Baie d'Audierne region is 070° (ref. 6); that of the migmatite belt to the east-south east is 110 – 120° (ref. 1). A gently curved arc can be envisaged, truncated by the late Variscan granites and mylonites of the South Armorican Shear Zone. Before the emplacement of these granites and the subsequent shear zone movement, the Baie d'Audierne group, which extends westwards to the edge of the submarine continental shelf⁹, would have extended to the east-north east, around the northern flank of the migmatite belt (Fig. 2). Whether or not the migmatite belt extended south of the present outcrop of the Baie d'Audierne group cannot be ascertained, as this position is occupied by a large, late Variscan granite (Pont l'Abbé). It is with this pre-late Variscan granite configuration in mind that the tectonic significance of the Baie d'Audierne group is discussed.

At, or even before, ~420 Myr ago, blueschist metamorphism was taking place to the south of a high temperature-medium

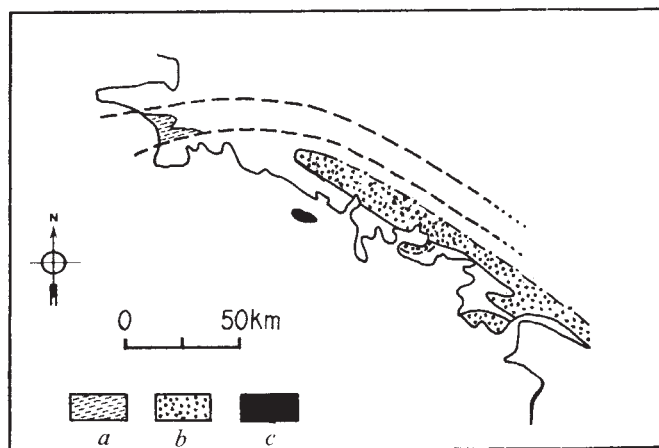


Fig. 2 Proposed configuration of rocks of Baie d'Audierne (a), Migmatite Belt (b) and Ile de Groix blueschist (c) types pre-emplacement of the late Variscan granites. The true location and lateral extent of (c) cannot be ascertained²⁴. Dashed lines indicate suggested boundaries.

pressure belt within which anatexis eventually occurred². Relict areas of older granitic material (Port Manech and Cap Sizun gneisses) within the western part of the high temperature zone^{1,6,17} suggest that the migmatite belt lay on, or immediately to the south of, continental crust (Fig. 3). The metamorphic belt therefore comprised an Andean type margin with a northward dipping subduction zone^{2,20}.

Metamorphism and deformation of this margin during the Ordo-Silurian (Ligerian) orogeny³ was related to subduction and the eventual arrival of northward moving continental material (Fig. 3), now lying immediately south of the blueschists (micaschists, granites and undifferentiated Palaeozoic sediments of the South Brittany continental shelf^{2,21}). In the Baie d'Audierne area, sometime between the initiation of subduction (~420 Myr) and 334 Myr ago, rifting occurred

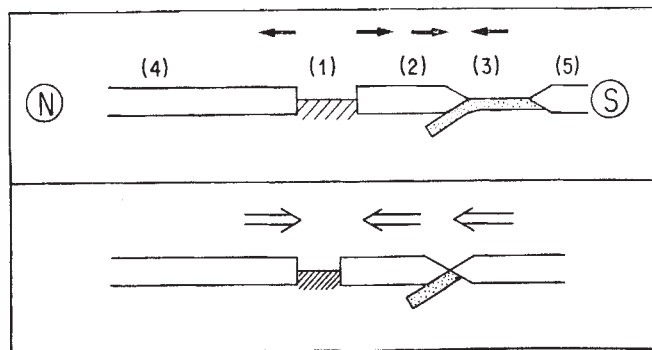


Fig. 3 Non-scaled diagram of postulated plate boundary relationships during the Palaeozoic, representing an hypothetical approximately north-south cross section across Fig. 2: 1, Baie d'Audierne type basic and ultrabasic rocks; 2, south Brittany migmatites; 3, Ile de Groix blueschists associated with oceanic crust; 4, northern continental block; 5, southern continental block. Post-420 Myr/pre-334 Myr plate motions (solid arrows) involve northward subduction at the southern margin of the migmatitic island arc (2), behind-arc extension (1) and northward movement of the southern continental block. Post-334 Myr, the behind-arc rift closes (open arrows, as the island arc + southern continent (2+5) collide with the northern continental block (4). (Adapted after Cogné²).

within the continental plate just to the north of the migmatite belt (Fig. 3). Into this rift, mantle derived gabbro and peridotite were intruded, representing the initial stages in the development of a marginal basin behind an island arc²², now deeply eroded and represented by the migmatite belt. (A mid-ocean ridge origin for these basic and ultrabasic rocks has also been suggested^{8,23}.)

At or post ~334 Myr ago, the migmatitic island arc, then in contact with the southern continental block (Fig. 3), was pushed northwards against the northern continental block, folding the behind arc rift zone and producing the Variscan deformation and metamorphism seen in the Baie d'Audierne. This continent plus island arc-continent collision corresponds to Cogné's intercontinental collision of the same age². In Cogné's model, however, the 'suture' representing this collision is marked by the South Armorican Shear Zone to the north (Fig. 1).

The south-armoricain metamorphic belt was situated therefore, at an active Andean type continental margin about 420 Myr ago, which after passing through a phase of back-arc rifting, progressed to a collision with a southern continent plus island arc and the northern continent, post 334 Myr ago.

I thank A. J. Barber, M. S. N. Carpenter and J. J. Peucat for helpful criticisms. A Chelsea College Research studentship is gratefully acknowledged.

S. K. HANMER

Centre Armorican d'Etude Structurale des Socles
Avenue du Général Leclerc
B.P. 25 A, 35031 Rennes Cédex, France

Received 4 May; accepted 19 September 1977.

- Cogné, J. *Mem. Serv. Carte Géol. Fr.* 370 (1960).
- Cogné, J. *Coll. Int. C.N.R.S. Géol. Himalaya* (in the press).
- Cogné, J., Jeannette, D. & Ruhland, M. *Bull. Serv. Carte Géol. Als. Lorraine* 41-95 (1966).
- Cogné, J. *Bull. Géol. Soc. Fr.* 4, 413-430 (1962).
- Cogné, J. in *Géologie de la France* (ed. Debelmas, J. 105-161 (Doin, Paris, 1974).
- Hanmer, S. K. thesis, Univ. London (1977).
- Cogné, J. *C.r. Acad. Sci. Paris*, 261, 3849-52 (1965).
- Peucat, J. J. thesis, Univ. Rennes (1973).
- Lefort, J. P. & Peucat, J. J. *C.r. Acad. Sci. Paris* 279, 635-637 (1974).
- Cogné, J. & Peucat, J. J. *C.r. Acad. Sci. Paris* 277, 2601-04, 1973.
- Vidal, Ph. *Bull. géol. Soc. Fr.* 15, 239-245 (1973).
- Audren, Cl. thesis, Univ. Rennes (1971).
- Hassenforder, B. thesis, Univ. Strasbourg (1970).
- Makanjuo, Z. A. & Howie, R. A. *Contr. Mineral. Petrol.* 35, 83-118 (1972).
- Carpenter, M. S. N. & Civetta, L. *Nature* 262, 276-277 (1976).
- Vidal, Ph. thesis, Univ. Rennes (1976).
- Peucat, J. J. & Cogné, J. *Nature* 268, 131-2 (1977).
- Audren, Cl. & Le Metour, J. *Bull. géol. Soc. Fr.* 18, 1041-49 (1976).
- Carpenter, M. S. N. thesis, Univ. Oxford (1976).
- Nicholas, A. *Nature* 236, 221-223 (1972).
- Lefort, J. P. thesis, Univ. Rennes (1976).
- Oxburgh, E. R. *Proc. geol. Ass.* 85, 299-357 (1974).
- Peucat, J. J. *Reun. Ann. Sci. Terre* 313 (1974).
- Audren, Cl. & Lefort, J. P. *Bull. géol. Soc. Fr.* (in the press).

Spike responses of 'non-spiking' visual interneurone

OUR understanding of information processing in nerve nets has been modified by the concept of graded signal transmission¹. Descriptions of non-spiking interneurons in insects²⁻⁸, and the demonstration of graded synaptic transmission⁹ have contributed to this development. In the fly visual nervous system, second and higher order interneurons are known, which apparently do not produce action potentials⁶⁻⁸. We show here that at least eight individually identifiable movement-sensitive cells, which have the characteristic properties of non-spiking interneurons¹ will generate spikes with imposed hyperpolarisation. Their graded mode of operation is due to maintained refractoriness. This applies selectively to neurones, which belong to either of two anatomically, and physiologically distinct classes. Other cell types in the same preparation generate spikes spontaneously.

The third visual neuropil of flies contains two conspicuous sets of directionally-selective movement-sensitive neurones^{9,10}—an array of 10 'vertical cells' (VS 1-VS 10) which respond to vertical movements in the ipsilateral visual field, and an array of three 'horizontal cells' (HS 1-HS 3) which respond to horizontal rotary movements in the visual field of the two eyes^{7,8}. Figure 2a-f shows some of the characteristic responses to visual stimuli of the neurone VS 1 depicted in Fig. 1. Its membrane potential in the dark is about -41 mV, modulated by sparse synaptic activity (Fig. 2a). Illumination of the receptive field causes a small depolarisation and a marked increase in 'noise' (Fig. 2b). When a striped pattern moves upwards in the receptive field, the cell is hyperpolarised, and the noise reduced (Fig. 2c). With downward movements it is depolarised, and the noise increased (Fig. 2d). Clockwise (Fig. 2e), or counter-clockwise (Fig. 2f) horizontal movements are ineffective. The movement responses are largest at a pattern speed of 2 periods per s, thus corresponding to the velocity-dependent optomotor behaviour of intact flies¹¹⁻¹³. It is possible to record this kind of electrical activity from any one of the VS- or HS-cells for more than 1 h, without ever

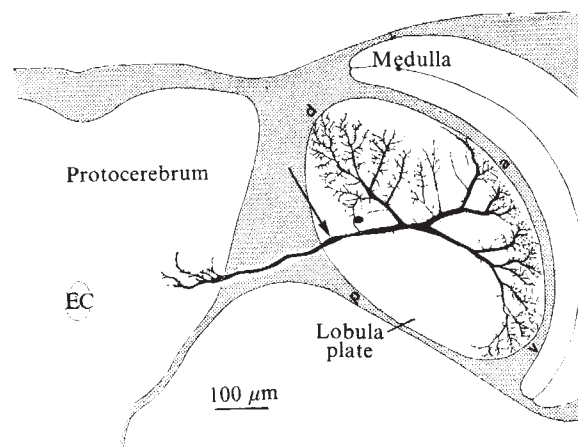


Fig. 1 Neurone VS1 of the blowfly *Calliphora erythrocephala*, as reconstructed after procion yellow injection, and serial sectioning. The right half of the fly's brain is shown from behind. Neuropil areas are white, tracts and perikaryon layers are shaded. The retinotopic projection of the ipsilateral visual field into the plane of the third visual neuropil (lobula plate) is indicated. VS1 belongs to a set of 10 large movement sensitive neurones, which occupy the caudal face of the neuropil. Their dendritic arborisations are in the lobula plate, the somata are in the caudal surface layer, and their axon terminals are close to the oesophageal canal (EC). Field axes: a-p, antero-posterior (horizontal); d-v, dorsoventral (vertical); arrow: site of penetration.

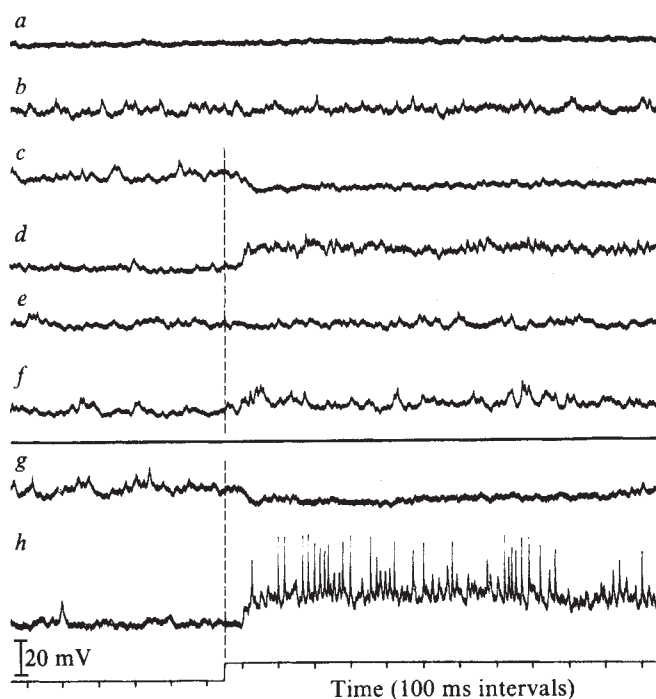


Fig. 2 Responses of neurone VS1 to visual stimulation. *a*, In darkness the membrane potential is $V = -41$ mV, and the average noise is 0.6 mV r.m.s. *b*, With constant illumination V shifts by +2.4 mV, but the noise increases threefold (1.8 mV r.m.s.). *c-h*, Responses to pattern movements. The pattern is illuminated throughout, and starts moving at the broken line. *c*, Upward movement hyperpolarises by -8.4 mV, and reduces the noise; *d*, downward movement depolarises by 11.0 mV, and increases the noise; *e*, clockwise, or *f*, counterclockwise horizontal movements give no significant responses. No spike is present in *a-f*. With steady hyperpolarisation, by injection of -2.4 nA, upward movement (*g*) still elicits a graded hyperpolarising response (-7.6 mV), but the effect of downward movement (*h*) is drastically different from (*d*): as well as an increased average depolarisation (+17 mV) fast spikes of about 30 mV amplitude are generated. These cannot be elicited by steady injection of depolarising current. Stimulus field: 45° lateral, 60° diameter; pattern: grating of 10° period; speed: 20° s^{-1} ; brightness: 50 cd m^{-2} , contrast: 80%.

observing a full-sized action potential. No spikes are elicited by various other kinds of visual stimuli (such as intense flashes, discrete moving objects) or by depolarising currents up to 10 nA.

VS- and HS-cells, like the neurone VS1 shown here, apparently have the specific properties of 'non-spiking' interneurons², namely: (1) small average membrane potential (-30 mV to -50 mV), (2) conspicuous noise-like potential fluctuations, (3) graded response to stimulation, (4) inability to generate action potentials with imposed depolarisation, and (5) graded transmission across chemical synapses of either polarity. This last property could not be tested here by simultaneously recording from postsynaptic neurones, but estimated length constants in VS and HS axons would allow graded potential changes to spread from the dendritic arborisation to the axon terminals. Any VS or HS cell will, however, generate spikes, if it is steadily hyperpolarised by current injection. The current required varies from less than -1 nA to about -5 nA and is inversely correlated with the membrane potential before current injection. The records in Fig. 2*g, h* are from the same neurone, VS1, as those in Fig. 2*a-f*, with the only difference that -2.4 nA are continuously injected through the recording microelectrode by means of a balanced bridge circuit. The cell responds to pattern movements as before, but Fig. 2*h* shows in addition fast spikes which often ride on top of small depolarising voltage fluctuations. These

spikes have a fast repolarising phase, their amplitude increases with current and may overshoot ground potential. The current-induced hyperpolarisation of the cell membrane does not exceed the equilibrium potential of the inhibitory synaptic ion channels: as Fig. 2*g* shows, there is still a hyperpolarising response. The current effect is reversible, and with steady injection, the generation of spikes may continue for more than 2 h. These findings were repeatedly confirmed for 6 of 10 VS cells, and for 2 of 3 HS cells. Figure 3, for instance, shows spikes from the neurone HS1 which are elicited by excitatory post-synaptic potentials (e.p.s.ps), but only if the cell is hyperpolarised. The same effects are expected from the remaining 6 neurones in the two classes, and are independent of the electrode-filling solution (5% Procion yellow M4RAN, 5% Procion rubine MX-B, isotonic potassium chloride).

The relationships between injected current, membrane potential, spike rate and spike amplitude, as studied more extensively in a VS neurone, are shown in Fig. 4. With increasing hyperpolarising current the spike rate rises rapidly from -1.5 nA to -3.0 nA, and declines again to zero between -3.0 nA and -5.5 nA, although the stimulation was stationary (constant light). The average membrane potential changes from -30 mV at 0 nA to -60 mV at -5.0 nA. The spike amplitude increases non-linearly with current, and overshoots ground potential if more than -3.0 nA are injected. The significant features of these spikes, namely their large amplitude, positive peak potential, and fast rate of repolarisation are not likely to arise from either chemical or electrical synaptic potentials, charging an otherwise electrically passive membrane.

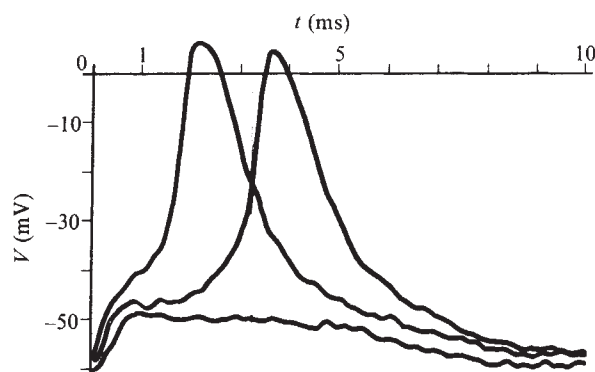


Fig. 3 Overshooting action potentials from neurone HS1, during steady hyperpolarisation with -1.0 nA. The spikes are elicited by large e.p.s.ps, evoked by horizontal pattern movement in the contralateral visual field. In one of the three superimposed traces no spike is generated and the e.p.s.p. is clearly visible. Records were taken at 5°C in order to spread the membrane processes in time, thereby circumventing the lowpass characteristic of the microelectrode ($240 \text{ M}\Omega$; 5% procion yellow M4RAN).

Rather, they demonstrate the existence of voltage-controlled conductance channels with kinetic properties appropriately tuned to produce regenerative action potentials. Accordingly Fig. 4 indicates that the current range of more than -5 nA corresponds to the resting state in a conventional neurone, -2 nA to -5 nA to an excited state, whereas between 0 and -2 nA the membrane is refractory. In a schematised notion of electrical excitability this means that most sodium channels are inactivated, and a large proportion of potassium channels are open. This state would be unstable unless the resulting outward potassium current is balanced by an inward current of equal strength. Either leakage or a constant and strong excitatory synaptic input could provide the inward current. Switching off the stimulating light reduces the synaptic 'noise' (Fig. 2*a*) without significant increase in membrane potential. The more

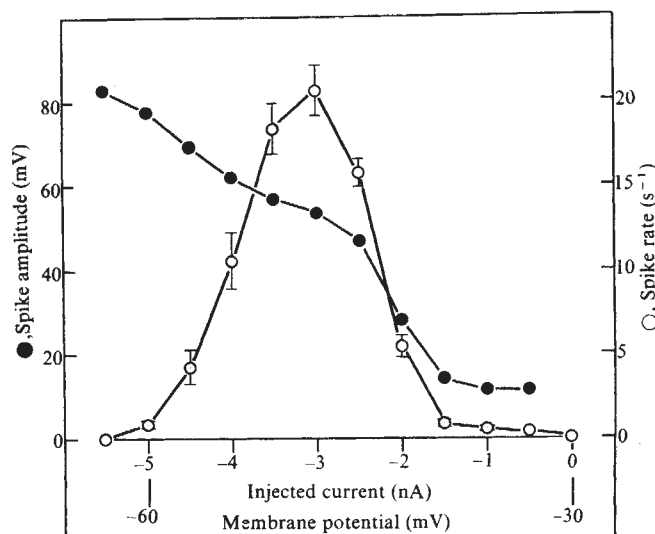


Fig. 4 Spike rate (○) and spike amplitude (●) as functions of injected current in a VS neurone of *Calliphora*. Weak stationary excitation is provided by constant illumination (Fig. 2b). With increasing hyperpolarisation of the cell membrane, the spike rate ($\pm$ s.e.m.) rises and falls steeply. At -5.0 nA the membrane potential is -60 mV, and the spike amplitude of 80 mV overshoots ground potential. The dependence of spike rate on membrane polarisation, the overshoot, and the fast repolarising phase (Fig. 3) characterise regenerative action potentials.

probable cause of the depolarised state is, therefore, a steady inward leakage current.

An interesting consequence with respect to the 'noisy' appearance of the membrane potential in the non-spiking state arises from the rapid breakdown of spike rate and spike amplitude between -2.5 and -1.5 nA in Fig. 4. In this poorly-defined state of the regenerative mechanism small fluctuations of synaptic or leakage currents will be amplified into comparatively large potential fluctuations. Noise measurements in this experiment yielded 0.78 mV r.m.s. at 0 nA, 0.75 mV r.m.s. at -5.5 nA, but 2.8 mV r.m.s. at -2.0 nA, where recognisable spikes contribute less than 0.3 mV r.m.s. to the overall noise.

These results show that VS and HS neurones have electrically excitable membranes. Their graded mode of operation is due to refractoriness of the regenerative membrane mechanism. Consequently neither natural nor imposed depolarisation will elicit spikes. The conspicuous membrane potential 'noise', which characterises the non-spiking state, is mainly caused by abortive membrane excitation. It is not clear why these cells in particular should be refractory, whereas other cells generate spikes. The advantage of permanent refractoriness is questionable; moreover, it has been shown in *Limulus* that electrical excitability can be differentially damaged in closely adjacent cells¹⁴. The 'non-spiking' state may therefore not represent the natural condition of information processing in these neurones.

ROLAND HENGSTENBERG

Max-Planck-Institut f. biologische Kybernetik,
D-7400 Tübingen, FRG

Received 3 May; accepted 10 August 1977.

- Schmitt, F. O., Dev, P. & Smith, B. H. *Science* **193**, 114–120 (1976).
- Pearson, K. G. & Fuortner, C. R. *J. Neurophysiol.* **38**, 33–52 (1975).
- Burrows, M. & Siegler, M. V. *Nature* **262**, 222–224 (1976).
- Chappell, R. L. & Dowling, J. E. *J. gen. Physiol.* **60**, 121–147 (1972).
- Laughlin, S. B. *J. comp. Physiol.* **112**, 199–212 (1976).
- Zettler, F. & Järvilehto, M. *J. comp. Physiol.* **85**, 89–104 (1973).
- Dvorak, D. R., Bishop, L. G. & Eckert, H. E. *J. comp. Physiol.* **100**, 5–23 (1975).
- Hausen, K. *Z. Naturforsch.* **31c**, 629–633 (1976); thesis, Univ. Tübingen (1976).
- Pierantoni, R. *Cell. Tiss. Res.* **171**, 101–122 (1976).
- Strausfeld, N. *J. Atlas of an Insect Brain*, Table 7.29 (Springer, Heidelberg, 1976).
- Götz, K. G. *Kybernetik* **2**, 77–92 (1964); **4**, 199–208 (1968).
- Götz, K. G. & Wenking, H. *J. comp. Physiol.* **85**, 235–266 (1973).
- Reichardt, W. & Poggio, T. *Q. Rev. Biophys.* **9**, 311–438 (1976).
- Barlow, R. B. & Kaplan, E. J. *J. gen. Physiol.* **69**, 203–220 (1977).

Evolutionary adjustment of developmental time in mixed populations of flour beetles

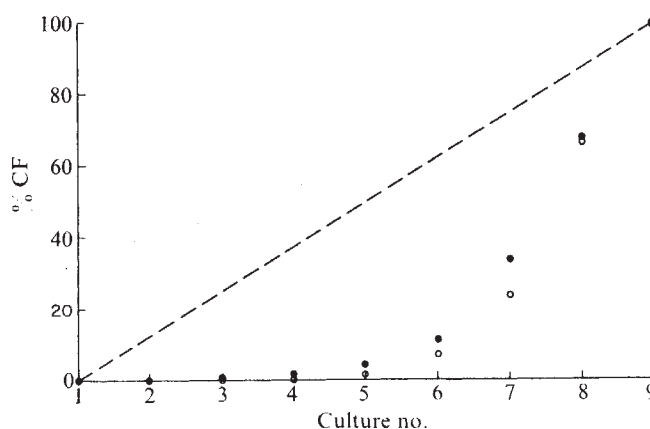
IN dense, synchronised laboratory populations of the flour beetle *Tribolium castaneum* (referred to here as CS), there is a substantial amount of cannibalism of the earliest pupating individuals by large larvae, thus creating the possibility of strong selection pressure against fast development^{1–3}. An opportunity to investigate this possibility was provided by the existence of a set of mixed-species cultures that had been maintained for several years. These cultures contain CS and *T. confusum* (CF) in varying proportions. Since developmental time of CF is from 1–2 d longer than CS⁴, the null hypothesis of no differences in developmental time of CS from the various cultures could be tested against the alternative that developmental time of CS would increase as the proportion of slower developing CF in the culture increases. I report here that developmental times do in fact adjust in mixed populations of these flour beetles.

As part of a study of changes in competitive ability and other fitness components in populations of CS exposed to differing proportions of intra- and interspecific 'competition', two replicate series of nine cultures were established in 1968 using beetles from new, genetically heterogeneous stocks of CS and CF. The cultures are maintained at 29°C , 50–70% relative humidity in half-pint milk bottles containing about 30 g of whole wheat flour enriched with 5% brewer's yeast. At the beginning of each transfer, or cycle of selection, the nine cultures receive 200 adults, distributed among the two species as follows: culture number (1), 200 CS/0 CF; (2), 175 CS/25 CF; (3), 150 CS/50 CF; (4), 125 CS/75 CF; (5), 100 CS/100 CF; (6), 75 CS/125 CF; (7), 50 CS/150 CF; (8), 25 CS/175 CF; (9), 0 CS/200 CF. The initial adults are discarded 3 weeks after being placed in the cultures, thus insuring a minimum of one generation in each transfer. About 7 weeks later, the cultures are censused and the initial composition is re-established.

In these conditions CF is clearly inferior to CS; Fig. 1 shows that the percentage of CF recovered is always lower than the percentage used to initiate the cultures. In fact, only cultures 8 and 9 (often 7 and occasionally 6) produce enough CF adults to re-establish the initial number. The CF beetles used to reinitiate the deficient cultures come from cultures 8 or 9.

Developmental time was measured in the summer of 1973 on the progeny of beetles taken from the 28th transfer using standard procedures⁴. For each culture in each of the two series, two replicate vials were set up containing 150 eggs 0–24 h of age in 6 g of flour–yeast medium. Numbers of pupae were recorded by sex at daily intervals starting 18 d after egg laying.

Fig. 1 % CF recovered from cultures 1–9 of the two series, compared to the % used to initiate the cultures (dotted line) for the first three transfers, or cycles of selection. Open and closed circles represent the two replicate series of cultures.



Since developmental time is not normally distributed⁴, analyses were done on vial means, which were based on from 98 to 137 individual observations.

Developmental time data are shown in Fig. 2. Cultures 8 are obviously very different from cultures 1–7. Ignoring cultures 8, a factorial analysis of variance was carried out using the vial means from cultures 1–7. Main effects were series (2), cultures (7) and sexes (2). The only significant main effect was cultures ($F = 9.88$, $P < 0.001$). The culture $\times$ series interaction was significant at the 0.05 level; this effect is a real one as will be shown below. Although differences between males and females have been found in some studies⁴, they were not evident here. Sexes were pooled for subsequent analyses.

Regression of developmental time on culture number for cultures 1–7 was carried out separately for each series. Linear regression was a significant source of variation for both series, and in one case there were also significant deviations from regression. The test for equality of slopes of the two regression lines showed that the slopes were not different, so a common line is plotted in Fig. 2. There is clear evidence for an increase in developmental time as a function of culture number.

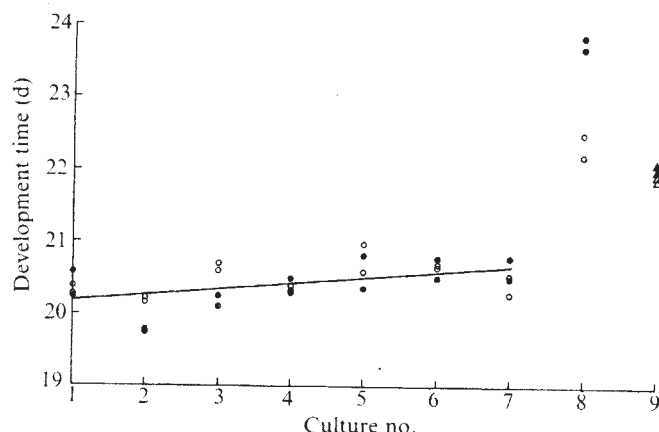


Fig. 2 Developmental time (from egg to pupa) for CS from cultures 1–8 and CF from culture 9. Data points are means of 98–137 individuals. Open and closed circles represent the two replicate series of cultures. The line from cultures 1–7 is a common regression line for the two series. The regression equation is $y = 20.12 + 0.08x$.

Turning now to cultures 8; developmental times changed dramatically to the point where they are even slower than the CF controls in cultures 9. It might be argued that this is a result of inbreeding rather than natural selection, since cultures 8 are reinitiated with only 25 adults at each transfer and developmental time is known to be subject to severe inbreeding depression in CS⁵. I argue here, however, that the change is an adaptive one.

First, inbreeding depression also leads to a great increase in variability of developmental time⁵, as does artificial selection for slow development⁴. Coefficients of variation for developmental time in controls (cultures 1) and cultures 8 were 5.84 and 7.78 for series A, and 8.68 and 8.33 for series B; clearly in series B at least the coefficient of variation is not larger in culture 8. This argues against inbreeding depression.

Second, fecundity is known to be also subject to severe inbreeding depression in this species⁵. Fecundity data, obtained in separate egg cannibalism experiments, will be published elsewhere. In one of the two series, the fecundity of females from culture 8 was not lower than that of the other populations, and in the other series there was a linear decrease in fecundity with culture number. These results do not support the hypothesis of inbreeding depression.

Finally, note that cultures 8 are the only ones where CF always replaces itself; 70–90% of the surviving adults at the time of census are CF. In addition, there is a marked similarity

between Figs 1 and 2. Pearson product-moment correlation coefficients were calculated for developmental time versus percentage CF for both the 1st and the 28th transfers (Table 1). The correlations for transfer 1 are both significant, in agreement with the hypothesis. Most remarkable, however, are the extremely high correlations at transfer 28. In the two series, 94 and 99% of the between-culture variation in developmental time can be accounted for by differences in the percentage CF among surviving adults.

Table 1 Coefficients of correlation (r) and determination (r^2) between mean developmental time in CS and competitive ability

Series		Transfer 1	Transfer 28
A	r	0.76*	0.97†
	r^2	0.57	0.94
B	r	0.89†	0.99†
	r^2	0.79	0.99

Competitive ability was measured as % CF among surviving adults.

* $0.05 > P > 0.01$.

† $P < 0.01$.

The hypothesis that developmental time of CS should increase as the proportion of slower developing CF in the cultures increases is supported by my data. The proposed mechanism of selection for slow development is cannibalism of early pupating individuals by large larvae. This mechanism was first suggested by Sokal and Sonleitner¹ as an explanation for increased developmental time in experimental CS populations. Using an artificial mixture of genotypes, I have shown that in dense cultures of CS there was strong selection against early pupation³. With increasing culture number in the present experiments there is a linear increase in the proportion of CF adults used to initiate the culture, but a distinctly non-linear increase in the proportion of CF among the survivors (Fig. 1). Developmental times of CS from these cultures form a very similar non-linear curve (Fig. 2). The high correlations between developmental time and percentage CF in the cultures at the time they were first established might conceivably be spurious ones involving some other, unknown, selective agent. But the extremely high correlations at transfer 28 are clearly suggestive of an adaptive response.

I thank Russel Riddle and Catherine Samollow for assistance, and John Wiens for comments on the manuscript. The research was supported by the USNSF.

PETER S. DAWSON

Department of Zoology,
Oregon State University,
Corvallis, Oregon 97331

Received 8 July; accepted 19 September 1977.

1. Sokal, R. R. & Sonleitner, F. J. *Ecol. Monogr.* **38**, 345–379 (1968).
2. Wool, D. *Res. Pop. Ecol.* **11**, 40–44 (1969).
3. Dawson, P. S. *Genetics* **80**, 773–783 (1975).
4. Dawson, P. S. *Genetics* **51**, 873–885 (1965).
5. Dawson, P. S. *Evolution* **22**, 217–227 (1968).

Bismuth staining of Golgi complex is a characteristic arthropod feature lacking in *Peripatus*

THE traditional view for the origin of arthropods is that an exoskeleton has evolved once and the three major modern groups, Crustacea (for example, crabs), Chelicerata (spiders) and Uniramia (insects) had a common ancestor in the pre-cambrian. This ancestor was itself related to primitive Onychophora (present day representative, *Peripatus*) and ultimately to predecessors of the annelids^{1–4}. But, arthropods may have had two, three or several separate evolutionary origins if the external features thought to unite the groups have evolved in response to functionally similar requirements^{3,5}. Studies on

trilobite skeleton, for example, suggest two independent ancestral groups whose members were not themselves arthropods. These gave rise on the one hand to the Uniramia and on the other to the chelicerates which were linked to Crustacea through ancestral trilobites⁶. The discovery of a cell component characteristic of all cell types in insects⁷ prompted a systematic study of its distribution, to see if it could be of help in determining arthropod evolutionary relationships. We report here that a characteristic feature of arthropods, bismuth staining of the Golgi complex beads, is absent in *Peripatus*, thus supporting the established view of arthropod evolution.

The smooth surface of the rough endoplasmic reticulum that makes the forming face of the Golgi complex in all cell types in *Calpodes* (Lepidoptera, HesperIIDae), has particles arranged like beads on a circular string with about 10 beads per ring and a centre-to-centre spacing of 27 nm. Each particle has a core 11 nm in diameter that stains specifically with bismuth. The rough endoplasmic reticulum membrane distends through the bead rings to form the transition vesicles of the Golgi complex. Particles with the distribution described for *Calpodes* may be present in most Golgi complexes but unless they stain with bismuth they are difficult or impossible to resolve. In some vertebrate tissues, for example, particles with the appropriate distribution do not react with bismuth and can only just be resolved after uranyl staining⁸. These results suggested that the beads or functionally comparable structures may be a general feature of Golgi complexes in organisms and that the bismuth staining might have a meaningful but restricted phylogenetic distribution. We therefore surveyed a number of arthropods and their near allies to determine the distribution of the bismuth staining of beads.

Tissues from organisms in the phyla listed in Table 1 were fixed in glutaraldehyde, stained only in bismuth solution before embedding⁹ and sections viewed with a Philips 300 electron microscope at a magnification of 250,000 and kept within 400 nm of focus. A piece of *Calpodes* tissue was carried through with each test specimen to make sure that failure to stain was not caused by an undetected failure of the technique. The success of the bismuth staining could also be judged by the reaction with interchromatin and perichromatin granules present in most nuclei. In practice there was rarely any doubt whether stained beads were present on suitably prepared Golgi complexes. They were usually found within minutes of beginning an observation or not at all, even after hours or days of searching. It may be that modification of the staining procedure could stain beads unrecognised hitherto, but the variations in staining recorded here reflect a consistent difference not attributable to variations in procedure.

Table 1 Groups in which bismuth-stained Golgi complex beads are present (+) and absent (—)

Cnidaria		(<i>Condylactis</i> , <i>Hydra</i>)	—
Platyhelminthes		(<i>Bipalium</i> , <i>Pseudoceros</i> , <i>Planaria</i>)	—
Nematoda		(<i>Phocinema</i>)	—
Mollusca		(<i>Placopecten</i> , <i>Anodonta</i>)	—
Annelida		(<i>Spirorbis</i> , <i>Sabella</i> , <i>Tubifex</i>)	—
Onychophora		(<i>Epiperipatus</i>)	—
Tardigrada		(<i>Hypsibius</i>)	+
Arthropoda	Pycnogonida	Pantopoda (Undetermined)	+
	Chelicerata	Xiphosura (<i>Limulus</i>)	+
		Acarina (Parasitidae)	+
		Pseudoscorpionidea (Undetermined)	+
	Crustacea	Isopoda (Undetermined)	+
		Decapoda (<i>Orconectes</i>)	+
	Diplopoda	Polydesmoidea (Undetermined)	+
	Insecta	Orthoptera (<i>Locusta</i>)	+
		Hymenoptera (<i>Apis</i>)	+
		Coleoptera (<i>Tenebrio</i>)	+
		Lepidoptera (<i>Calpodes</i>)	+

Figure 1 shows some of the results. The characteristic size and spacing described for bismuth-stained GC beads in *Calpodes* is a constant feature throughout the arthropods. Table 1 shows the phylogenetic distribution. The beads are present in all three major arthropod groups and also in pycnogonids and tardigrades. They are absent from Onychophora (*Epiperipatus*),

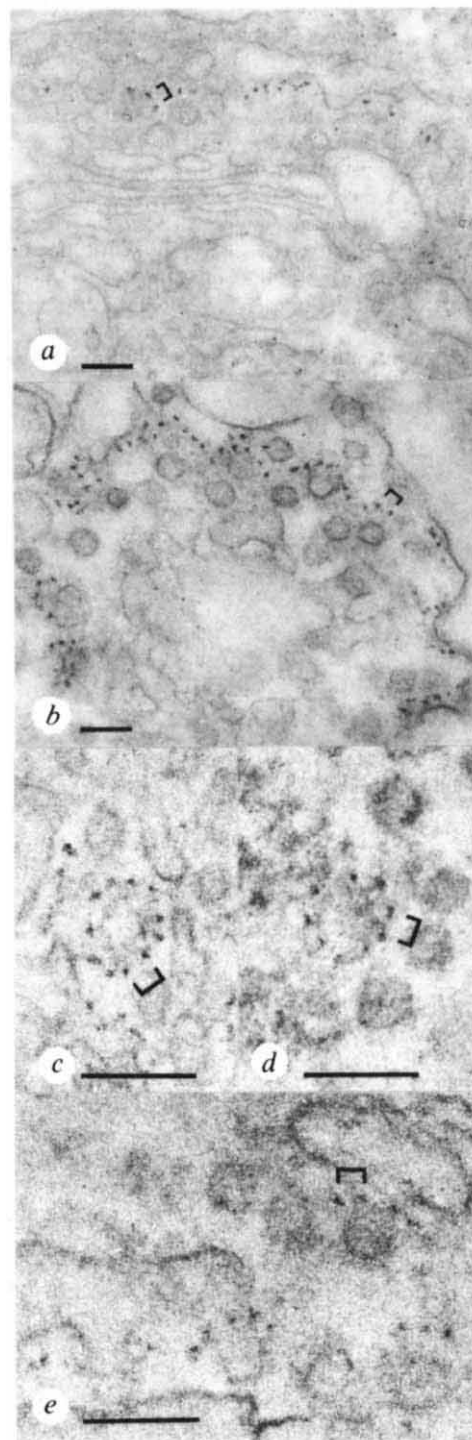


Fig. 1 Bismuth staining of the Golgi complex beads in the main arthropod lines of evolution. *a*, Uniramia, Insecta, *Calpodes*, Lepidoptera, HesperIIDae, larval epidermis. *b*, Crustacea, Decapoda, *Orconectes* x-organ. *c*, Chelicerata, Arachnida, Acarina, epidermis. *d*, Pycnogonida, Pantopoda, epidermis. *e*, Chelicerata, Merostomata, Xiphosura, *Limulus*, epidermis. The tissues were stained only by bismuth which reacts with particles about 10 nm in diameter arranged in rings upon the smooth face of the rough endoplasmic reticulum where transition vesicles of the Golgi complex arise. The square bracket marks the characteristic 27 nm spacing. Stained beads are absent from all tissues in *Epiperipatus*. Scale bar, 100 nm.

annelids, molluscs and flatworms. The inclusion of the tardigrades with arthropods agrees with work showing chemical similarities between the cuticles¹⁰. The distribution of beads thus favours a common origin for present-day arthropods and makes it unlikely that the Uniramia arose separately from a *Peripatus*-like ancestor.

This work was supported by a National Research Council of Canada grant (A6607).

M. LOCKE
P. HUIE

Cell Science Laboratories,
Department of Zoology,
The University of Western Ontario,
London, Ontario, N6A 5B7,
Canada

Received 5 August; accepted 16 September 1977.

1. Clarke, K. U. *The Biology of the Arthropoda* (eds Barrington, E. J. W. & Willis, A. J.) 172 (Edward Arnold, London, 1973).
2. Imms, A. D. *Outlines of Entomology* 167–170 (Methuen, London, 1942).
3. Manton, S. M. *Zool. J. Linn. Soc.* 51, 380 (1972).
4. Meglitsch, P. A. *Invertebrate Zoology* 668 (Oxford University Press, New York, 1967).
5. Tiegs, O. W. & Manton, S. M. *Biol. Rev.* 33, 255–337 (1958).
6. Cisne, J. L. *Science* 186, 13–18 (1974).
7. Locke, M. & Huie, P. *J. Cell Biol.* 70, 384–394 (1976).
8. Locke, M. & Huie, P. *Tissue Cell* 8, 739–743 (1976).
9. Locke, M. & Huie, P. *Tissue Cell* 9, 347–371 (1977).
10. Jeuniaux, C., *Cah. Biol. mar.* 16, 597–612 (1975).

Induction of fibrin thrombi by monocytes

THE discovery that blood clots in the absence of its cellular components has suggested that coagulation is independent of blood cells¹. This view has been modified by the demonstration of procoagulant factors in the platelets² and anticoagulant heparin in basophils³. Procoagulant activity has also been found in extracts of leukocytes^{4,5}, showing that coagulation is a cell-regulated phenomenon. We report here the direct observation of focal clotting induced by specifically activated monocytes.

Orientated fibrin micro-clots can be produced by incubation of bacterial endotoxin, the white cell buffy coat and heparinised plasma from sensitised patients⁶. The fibrin radiates in a stellate pattern from a centre of indeterminate nature (Fig. 1). We have analysed this central body.

When we used buffy coats rendered largely devoid of platelets by repeated centrifugal separation, the nucleus of the micro-thrombus appeared as a cell rather than an amorphous mass. Attempts to ascertain the responsible cell by using Ficoll, dextran and albumin separation techniques were unsuccessful because thrombi did not form. But direct centrifugal separation of untreated buffy coat specimens in plastic tubes⁷ revealed the upper portion of the buffy coat containing lymphocytes and monocytes to be associated with the induction of the greatest number of thrombi. The lower fraction containing the heavier polymorphonuclear leukocytes showed minimal activity. Rectangular capillary tube thrombi counts⁸ and differential counts of smears of comparable buffy coat fractions demonstrated a correlation between the number of thrombi and the number of monocytes or lymphocytes, but not the number of neutrophils, eosinophils or basophils.

The monocyte shows strong surface adherence in contrast to the lymphocyte, and we found that the cells responsible for the thrombi rapidly attached to glass. Buffy coat preparations left in contact with the glass capillary tube or on glass cover slips could be flushed or washed away after 30–60 min at 37 °C, leaving adherent cells which became centres for thrombus formation in the endotoxin-plasma mixture then applied. Furthermore, in preparations made on cover slip slides whereby the slip and slide were forced apart, the central cell had a remarkable set of pseudopods, as evidence of its strong attachment to both surfaces (Fig. 2). Blood left in contact with glass wool for 1 h and then removed failed to produce thrombi.

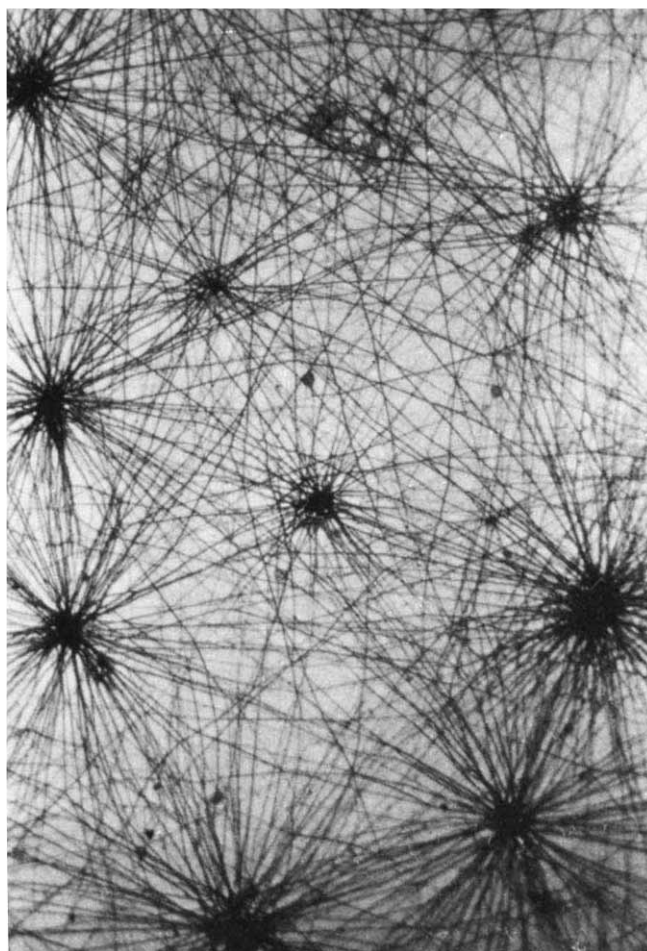


Fig. 1 Multiple stellate foci of fibrin induced by addition of lipopolysaccharide W (10^{-4}) to heparinised blood from patient with vasculitis. The cover slip preparation was incubated 18 h at 37 °C. Central bodies appear to be of indeterminate nature. (Wright-Giemsa.)

Conversely lymphocyte specimens (99% pure) obtained by passing blood through cotton⁹ produced no thrombi when exposed to endotoxin.

The cells responsible for the thrombi also proved to be phagocytic. The addition of latex particles⁹ ($0.011 \mu\text{m}$, 50λ of 10^{-4} in 1 ml blood) to the regular system led to their endocytosis by the central cell of many of the thrombi (Fig. 3).

Phase and phase interference microscopy with unsealed cover slip preparations showed that the centre of the thrombus was a large cell, generally without refractile granules, and with a reniform, pyknotic or vacuolated nucleus. The cell appearing to be a monocyte had fine filopodia at first. The earliest identifiable pericellular fibrin crystals appeared after 3.5 h. These were short and increased in number and length as time passed. The cell had grey cytoplasm and the nucleus became pyknotic, at time showing karyorrhexis and perinuclear vacuole formation. It sometimes ingested granules. The cell enlarged, became degenerate¹⁰, with cytoplasm of granular appearance, within 24 h (Fig. 3). At the end of this time, cell identification proved difficult in most cases because of the advanced degenerative changes. The spreading of the central cell could be accentuated markedly by placing a 30-g weight on the cover slip during incubation. Wright-Giemsa (5% acetic acid methanol fixation) or haematoxylin-eosin (Carnoy's fixative) stained the nucleus clearly and showed an eosinophilic cloud at the centre of the thrombus. The myeloperoxidase¹¹ stain (10% formalin alcohol) showed fine granules consistent with those seen in the monocyte, and neutral red staining showed supravital uptake by some of the central cells. The best stain proved to be the Papanicolaou¹² (95% ethyl alcohol fixation)

because it stained the cytoplasm as well as the nucleus. In the early phase the cytoplasm was green, but after about 6 h it was a distinctive pink orange and later yellow. This seemed to herald the death of the cell, and with this stain each thrombus could be readily spotted by the pinkish centre.

Immunofluorescent staining with fluorescein–isothiocyanate-conjugated anti-human immunoglobulins (Cappel) of cold-acetone-fixed cover slip specimens revealed the central cell stained with C_3 but not with IgG A M D or C_2 or C_2 . This is consistent with the fact that the monocyte has a receptor for C_3 (ref. 13).

The monocyte seems to have a clear cellular role in coagulation. Its ability to induce focal fibrin formation may be related to the observation that human monocytes exposed for 4 h at 37 °C to *Escherichia coli* lipopolysaccharide have a demonstrable coagulant effect in a one-stage coagulation assay¹⁴. Other non-specific effects of endotoxin include the release of pyrogen, as well as T-cell and colony stimulation¹⁵. This is in contrast to the specific phenomenon we have observed in patients with vasculitis and various immunological diseases. The fibrin deposition which we have observed around the monocyte may reflect complement activation through the alternative pathway.

As a result of these studies, the monocyte can be viewed as a clot cell, in Yin-Yang contrast to the heparin-containing basophil, an anticlot cell. Study of the monocyte in cell-mediated coagulation promises new information about thrombosis, gram-negative infections and various immunological diseases which involve fibrin deposition. So far the only clinical analogue of our laboratory micro-thrombi may be the stellate

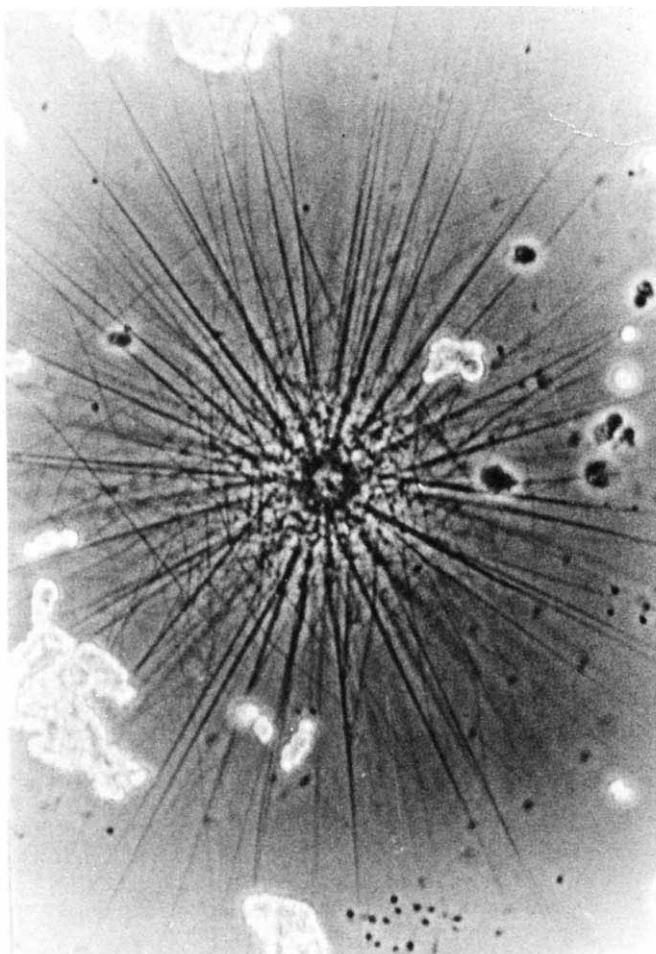


Fig. 3 Single stellate micro-thrombus forming in presence of latex particles. Refractile granules in central cell demonstrate it to be phagocytic. (Unstained, phase.)

fibrin formation observed in the bone marrow of patients with myelofibrosis¹⁶. The nature of this is not known but the bone marrow monocyte may play a central role.

We thank Marilyn Salscheider for technical assistance. This work was supported by grants from the Walter Annenberg Fund, the Thomas B. McCabe Fund and the Swedish Medical Research Council.

WALTER B. SHELLEY

Department of Dermatology,
University of Pennsylvania School of Medicine,
Philadelphia, Pennsylvania 19104

LENNART JUHLIN

Department of Dermatology,
University of Uppsala School of Medicine,
Uppsala, Sweden

Received 8 June; accepted 23 August 1977.

1. Tullis, J. L. *Clot* (Thomas, Springfield, Illinois, 1976).
2. Walsh, P. N. *Blood* 43, 597 (1974).
3. Parawaresch, M. R. *The Human Blood Basophil* (Springer, Berlin, 1976).
4. Niemetz, J., Muhlfelder, T., Chierago, M. E. & Troy, B. *Ann. N.Y. Acad. Sci.* 283, 208 (1977).
5. Rothberger, H., Zimmerman, T. S., Spiegelberg, H. L. & Vaughan, J. H. *J. clin. Invest.* 59, 549 (1977).
6. Juhlin, L. & Shelley, W. B. *Lancet* i, 21 (1977).
7. Shelley, W. B. *Am. J. clin. Path.* 39, 433 (1963).
8. Ling, N. R. & Kay, J. E. *Lymphocyte Stimulation* 2nd edn. (North Holland, Amsterdam, 1975).
9. Rothbart, P. H., Hendriks-Sturkenboom, I. & Ploem, J. S. *Blood* 48, 139 (1976).
10. Bessis, M. *Living Blood Cells and Their Ultrastructure* (Springer, New York, 1973).
11. Kaplow, L. S. *Blood* 26, 215 (1965).
12. Keebler, C. M. & Reagan, J. W. *A Manual of Cytotechnology* 4th edn (American Society of Clinical Pathologists, Chicago, 1975).
13. Seligmann, M. *Br. J. Haemat.* 31, Suppl. 1, 1 (1975).
14. Rivers, R. P. A., Hathaway, W. E. & Weston, W. L. *Br. J. Haemat.* 30, 311 (1975).
15. Territo, M. C. & Cline, M. J. *J. Immun.* 118, 187 (1977).
16. Nezelof, C., Soulier, J. P., Griscelli, C. & Royer, P. *Nouvelle Rev. Franc. d'Hematologie* 14, 641 (1974).

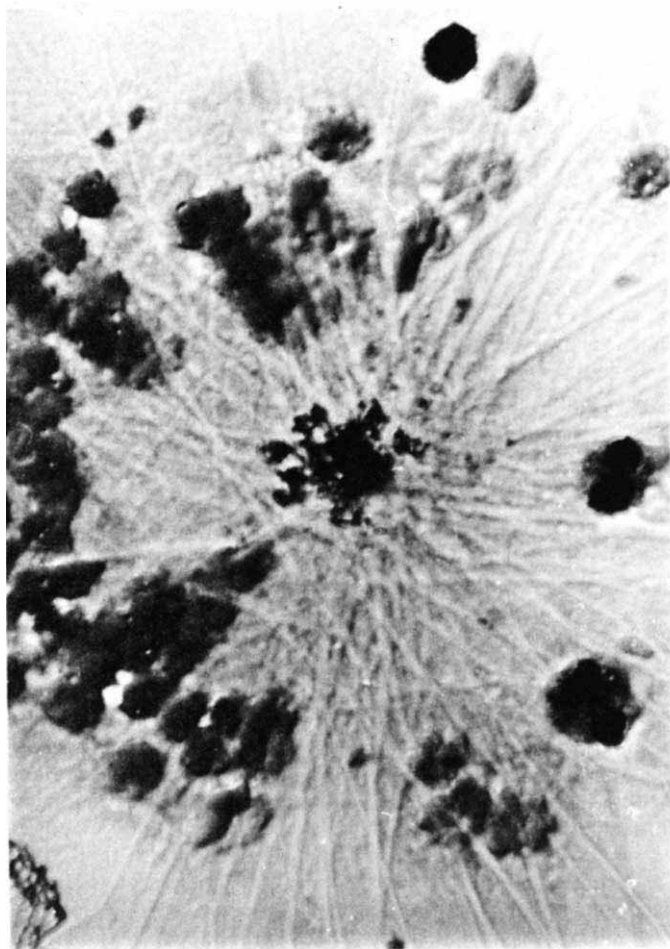


Fig. 2 Single stellate micro-thrombus forming in presence of buffy coat from which platelets have been removed. The unstained fibrin crystals radiate from monocyte, a central dark staining cell with pseudopods. Note lymphocyte at top two polymorphonuclear cells to right and peripheral red cell clusters. (Wright-Giemsa, phase.)

Morphological abnormalities in spermatozoa of man and great apes

HUMAN semen has long been known to differ from that of any other mammal in the high proportion of abnormal spermatozoa that it contains¹. But there has been almost no opportunity to compare human spermatozoa with those of man's closest living relatives, the chimpanzee (*Pan troglodytes*), the pygmy chimpanzee (*Pan paniscus*), the gorilla (*Gorilla gorilla*) and the orangutan (*Pongo pygmaeus*); these are all now endangered species, and the few animals in zoos are not available for experimentation. Recently, workers at the Yerkes Regional Primate Research Center, where all four species are maintained, have perfected a technique for obtaining semen from these rare animals by electroejaculation². Preliminary studies have suggested that marked morphological similarities exist only between the spermatozoa of the gorilla and man³. We examined these similarities and differences in more detail, in the hope that spermatozoal morphology might provide some taxonomic clues about our affinities to the great apes.

Measurements of total dry mass were made with an integrated interferometer (Vickers M-86) for a random selection of 50–100 spermatozoa from each individual in each of the five species. In man and gorilla the dry mass distribution was bimodal with two distinct non overlapping peaks, whereas in the chimpanzees and the orangutan it was unimodal (Fig. 2). When arbitrary units of dry mass were used, the mean value of the major peak in man was 730 compared with 960 in the gorilla. In both species the minor peak represented a dry mass approximately twice that of the major peak. These results are similar to those reported for Feulgen-stained human spermatozoa⁴, and suggest that the two groups represent haploid and diploid spermatozoa. The percentage of diploid spermatozoa in man and gorilla was estimated using a larger sample; 21 diploids were found in 2,000 human spermatozoa (1.05%) and 37 in 2,408 gorilla spermatozoa (1.53%). This difference was not statistically significant ($P > 0.10$). Human spermatozoa had the lowest mean haploid dry mass, and gorilla spermatozoa the highest (Table 2). By extracting DNA with trichloroacetic acid and remeasuring dry mass it could be shown that total dry mass was proportional to DNA content in *Homo sapiens*, *Pan troglodytes*, *Gorilla gorilla* and *Pongo pygmaeus* (Fig. 3). Statistical analysis (Table 2) revealed no signi-

Table 1 Morphology of spermatozoa of man and the great apes

		<i>Homo sapiens</i> (4 individuals)	<i>Pan troglodytes</i> (3 individuals)	<i>Pan paniscus</i> (1 individual)	<i>Gorilla gorilla</i> (2 individual)	<i>Pongo pygmaeus</i> (2 individuals)
Normal	Modal	73.0	95.5	98.0	71.0	98.5
Abnormal	Large head	2.0	0.3	—	2.3	—
	Small head	0.5	—	0.5	0.8	—
	Tapered head	0.3	0.2	—	1.0	—
	Dense-staining	0.6	0.2	0.5	4.5	—
	Vacuolated head	2.4	3.0	1.0	1.3	0.3
	Irregularly shaped	18.4	0.2	—	16.0	0.8
	Multiple heads	0.5	—	—	0.5	—
	Abnormal midpiece	—	0.3	—	0.5	—
	Cytoplasmic droplets	—	—	—	0.3	—
	Immature cells	2.3	0.3	—	2.0	0.5
	Total	100.0	100.0	100.0	100.2	100.1

Semen samples were obtained from normal healthy individuals of proven fertility except for one *Pongo pygmaeus*. In man and in *Pan troglodytes* samples were obtained by masturbation. In *Pan paniscus*, *Pongo pygmaeus* and *Gorilla gorilla* samples were obtained by electroejaculation as described by Warner *et al.*². Semen samples were washed in 0.9% saline and fixed in methanol: acetic acid, 3:1 as described by Sumner⁴. Some of these slides were stained with Papanicolaou's stain and observed under the light microscope: 200 cells were scored per individual and all results of Table 1 are expressed as percentages.

Table 1 shows that there is a similar proportion of 'normal' (modal) and abnormal forms in the ejaculate of men and gorillas; morphologically, the 'normal' forms of the two species are indistinguishable (Fig. 1). Chimpanzee spermatozoa are very different in appearance to human or gorilla spermatozoa, although very uniform in shape; they seem to be identical in all aspects to the spermatozoa of the pygmy chimpanzee. The spermatozoa of the orangutan are morphologically distinct from all the other species, although once again extremely uniform in shape.

Significant differences in dry mass between individuals within a species, but marked differences between species. Although human spermatozoa are morphologically most similar to those of the gorilla, in terms of DNA content they resemble those of the chimpanzees. Gorilla spermatozoa showed the greatest intra-individual variance in haploid DNA content.

Diploid DNA estimates in the somatic cells of the great apes and man have revealed *Homo sapiens*, < *Gorilla gorilla*, < *Pan troglodytes*, < *Pongo pygmaeus*⁵ which does not correspond with our haploid estimates of *Homo sapiens* < *Pan troglodytes*,

Table 2 Comparisons of total dry mass of spermatozoa of the great apes and man

Species	Means	Difference	Approximate 95% confidence limits			
			Assumption A		Assumption B	
			Lower	Upper	Lower	Upper
Chimpanzee-human	810.0–742.3	+67.7	+21.8	+113.6	+41.8	+93.6
Orangutan-human	871.6–742.3	+129.3	+73.0	+185.6	+97.6	+161.0
Gorilla-human	965.2–742.3	+222.9	+166.4	+279.5	+190.7	+255.2
Orangutan-chimpanzee	871.6–810.0	+61.6	+5.3	+117.9	+29.9	+93.3
Gorilla-chimpanzee	965.2–810.0	+155.2	+98.7	+211.8	+123.0	+187.5
Gorilla-orangutan	965.2–871.6	+93.6	+28.5	+158.9	+56.6	+130.7

The component of variance between individuals within a species is taken to be either 2,000 a.u.² (assumption A) or 1,000 a.u.² (assumption B).

Anovar of haploid dry-mass revealed (1) no significant differences between preparations from the same individual, (2) that the intra-individual variance components for man, chimpanzee and orangutan did not differ significantly (pooled estimate = 4,949.3 with 485 d.f.), but were significantly less than for gorilla (11,606.7 with 197 d.f.). Confidence intervals were constructed using these estimates and taking values of the inter-individual intra-species component to be either 2,000 a.u.² (assumption A) or 1,000 a.u.² (assumption B). The assumptions correspond to situation in which one individual in 100 can be expected to differ from its species mean by the equivalent of approximately two (A) or one (B) medium-sized chromosomes respectively. The implied degree of aneuploidy suggests that these are overestimates, as are the widths of the corresponding confidence intervals.



Fig. 1 The spermatozoa of *Pan paniscus* (a); *Pan troglodytes* (b); *Pongo pygmaeus* (c); *Homo sapiens* (d); and *Gorilla gorilla* (e) photographed under Nomarski interference. Scale bar, 10 μ m.

< *Pongo pygmaeus*, < *Gorilla gorilla*. Man has a diploid chromosome number of 46 compared with 48 in all great apes, and less constitutive heterochromatin and brilliant fluorescent chromatin than the chimpanzee and the gorilla⁶. The overall amount of satellite DNA sequences in man seems to be lower than in the great apes⁷, and presumably all these factors contribute to the low DNA content of human cells.

It has been suggested⁸ that the pleiomorphism of human spermatozoa may be a consequence of testicular damage caused by clothing-induced hyperthermia. The fact that the spermatozoa of gorillas living in outdoor cages are just as pleomorphic, and

show an even greater variability in haploid DNA content, tends to refute this argument. Furthermore, the great variability in morphology of human spermatozoa, in contrast to the lack of variability in haploid DNA content, means that morphologically abnormal human spermatozoa are not necessarily genetically defective. The percentage of morphologically abnormal forms is increased in infertile men, however, sometimes reaching very high proportions in cases of abnormal chromosome constitution⁹. It

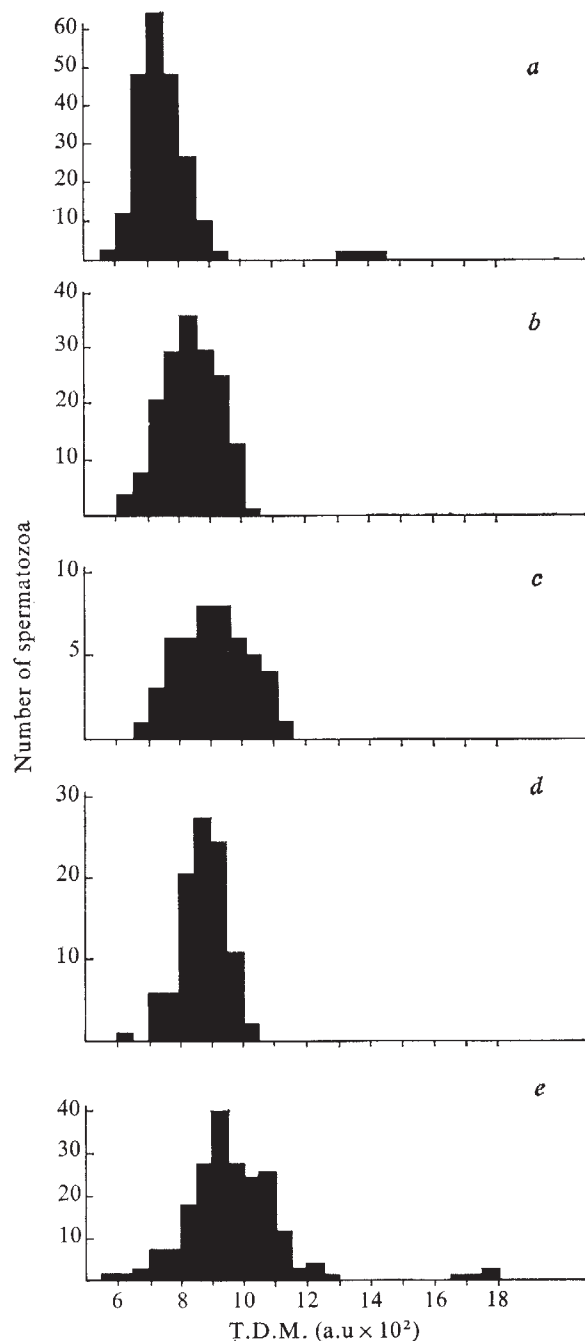


Fig. 2 The estimations of total dry mass (TDM) were obtained with a Vickers M-86 integrated microinterferometer. This instrument measures the optical path difference (o.p.d. = refractive index $\times$ thickness) in arb. units (a.u.). The area of the object (the sperm head only) was selected using an electronic masking system. Measurements were obtained with a X 75 n.a. 1.1 water immersion objective and preparations were measured while immersed in distilled water. All semen samples were treated as specified in the legend of Table 1, except that measurements were done on unstained preparations. 50 cells were measured per individual using two slides (25 measurements in each), except in the gorillas (100 cells; 50 measurements per slide). a, *Homo sapiens*; b, *Pan troglodytes*; c, *Pan paniscus*; d, *Pongo pygmaeus*; e, *Gorilla gorilla*.

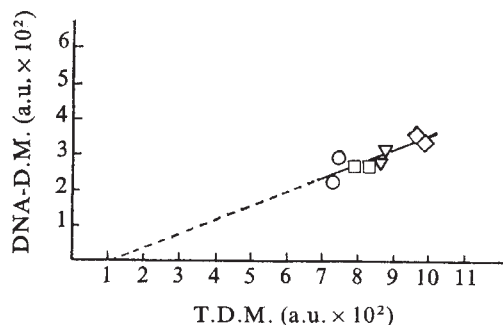


Fig. 3 To test whether TDM is proportional to DNA content, 10 spermatozoa from each of two individuals from each species (except *Pan paniscus*) were measured for TDM. After treatment with 5% trichloroacetic acid at 90 °C to remove DNA, the same spermatozoa were relocated and measured for dry mass after extraction (DMAE). Slides were then stained with Feulgen and the same spermatozoa measured for residual DNA with the M-86 instrument in the microdensitometric mode. Only if no DNA was detected were the DMAE values considered valid. The minimum extraction time needed to remove all DNA was 15 min in man and 30 min in the great apes. The difference between TDM and DMAE represents the dry mass of extracted DNA (DNA-DM). The figure shows mean values of DNA-DM and TDM for each individual, and the estimated least-squares regression line. The latter passes close to the origin, suggesting that proportional changes in TDM are identical to those for DNA-DM. (The line should not of course be interpreted as a valid extrapolation beyond the range of the plotted points.) ○, *Homo sapiens*; □, *Pan troglodytes*; ▽, *Pongo pygmaeus*; ◇, *Gorilla gorilla*.

has even been suggested that fertilisation of a defective egg by diploid spermatozoa will give rise to a hydatidiform mole¹⁰, and that 40% of triploid embryos are a result of fertilisation of a normal egg by a diploid spermatozoon¹¹.

Spermatozoal morphology has been shown to be a useful taxonomic guide in assessing relationships between other species¹². There is already evidence to suggest that man and gorilla are more closely related than any of the other great apes, as judged by their chromosomal karyotypes^{13,14}. Our data on the morphology of their spermatozoa lend further support to this conclusion.

H. N. SEUANEZ*
A. D. CAROTHERS

MRC Clinical and Population Cytogenetics Unit,
Western General Hospital,
Edinburgh, UK

D. E. MARTIN

Yerkes Regional Primate Center,
Atlanta, Georgia

R. V. SHORT

MRC Unit of Reproductive Biology,
2 Forrest Road, Edinburgh, UK

Received 11 July; accepted 23 September 1977.

*To whom reprint requests should be addressed.

- Bedford, M. *Contr. Primat.* 3, 97-139 (Karger, Basel 1974).
- Warner, H., Martin, D. E. & Keeling, M. E. *Ann. biomed. Engng* 3, 419-432 (1974).
- Martin, D. E., Gould, K. G. & Warner, H. *J. hum. Evol.* 4, 287-292 (1975).
- Sumner, A. *Nature* 231, 49 (1971).
- Manfredi Romanini, M. G. *J. hum. Evol.* 1, 23-40 (1972).
- Conf. Suppl. Standardization in Human Cytogenetics. Birth Defects Original Article Series I, 9* (The National Foundation, New York, 1975).
- Gosden, J., Mitchell, A., Seuanes, H. & Gosden, C. *Chromosoma* 63, 253-271 (1977).
- Bedford, M., Bent, M. J. & Calvin, H. *J. reprod. Fert.* 33, 19-29 (1973).
- Chandley, A. *et al. Ann. hum. Genet.* 39, 231-254 (1975).
- Kajii, T. & Ohama, K. *Nature* 268, 633-634 (1977).
- Beatty, R. A. *Physiology and Genetics of Reproduction Part A* (ed. Coutinho, A. M. & Fuchs, F.) (Plenum, New York, 1974).
- Baccetti, B. & Afzelius, B. A. *Monographs in Developmental Biology* (Karger, Basel, 1976).
- Seuanes, H., Robinson, J., Martin, D. E. & Short, R. V. *Cytogenet. Cell. Genet.* 17, 317-326 (1976).
- Seuanes, H. *thesis, Univ. Edinburgh* (1977).

Continuous growth and differentiation of human myeloid leukaemic cells in suspension culture

ATTEMPTS to develop long-term suspension cultures of human myeloid leukaemic cells have met with limited success. Lymphoblastoid lines carrying the Epstein-Barr virus genome occasionally arise during such attempts but these lymphoid cells originate from contaminating B lymphocytes and not from the leukaemic myeloid cells¹. A line established from the pleural fluid of a patient with chronic myeloid leukaemia in blast crisis² (designated K-562) has no B-cell or T-cell markers³⁻⁴ and does not seem to be of lymphoid origin⁴. Its lack of morphological and histochemical differentiation²⁻⁴, however, makes it difficult to determine whether these cells are derived from myeloblasts or more primitive stem cells⁴. Another less documented cell line (8261) derived from the peripheral blood of a patient with acute myelogenous leukaemia showed apparent morphological and functional differentiation in agar in the presence of a feeder layer of peripheral blood leukocytes but did not differentiate in suspension culture⁵. Our laboratory previously reported that cultures of differentiating myeloid leukaemic cells can be maintained for several months in suspension culture but only when enriched with conditioned media (CM) from certain monolayer fibroblastic cultures of first trimester whole human embryos (ref. 6 and Ruscetti *et al.* in preparation). We describe here for the first time the derivation from myeloid leukaemic cells of a leukocyte culture that by morphological and histochemical criteria clearly and persistently differentiates along the myeloid series without an exogenous source of conditioned medium.

Peripheral blood leukocytes (designated HL-60) from an adult female with acute promyelocytic leukaemia (clinical details of the case will be published elsewhere) were seeded in T30 plastic flasks (Falcon) at 1.25×10^6 cells per flask in 5 ml of RPMI-1640 (GIBCO) medium supplemented with 15% heat-inactivated foetal calf serum (Flow Labs) and gentamicin $50 \mu\text{g ml}^{-1}$. The flasks were incubated at 37 °C in a humidified 5% CO_2 atmosphere. A number of the

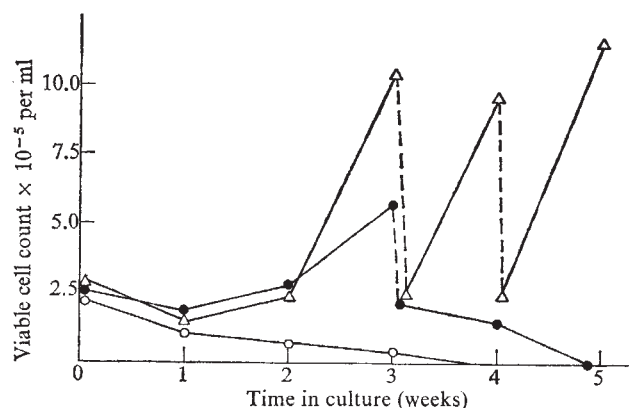


Fig. 1 Initial growth of HL-60 leukaemic leukocytes in suspension culture. After seeding the fresh peripheral blood leukocytes at 2.5×10^5 cells per ml, weekly haemocytometer counts were done on aliquots of cell suspensions grown in the presence or absence of conditioned media derived from human embryonic cultures. The dotted lines indicate a culture split. During these early passages, cell counts were done only at fixed intervals and hence do not necessarily indicate the highest count during each passage. TVL is a fibroblast culture of a 12-week human embryo lung. DHL is a fibroblastic culture of a 14-week human embryo lung. ●, TVL-conditioned media; △, DHL-conditioned media; ○, no conditioned media.

Table 1 Degree of differentiation of fresh uncultured HL-60 cells and cultured cells grown in the presence or absence of DHL-conditioned medium

	Blasts	Promyelo- cytes	Myelo- cytes	Segs* †bands †metas	% Per- oxidase positive	% Sudan black positive	Mono- cytes	Lympho- cytes	Eosin- ophils	Baso- phils	Nucleated RBCs
Fresh HL-60	25	49	3	1	55	NT	6	15	0	0	1
Cultured HL-60											
Day 2+CM	19	69	7	4	97	95	1	0	0	0	0
Day 2-CM	23	65	4	6	98	96	2	0	0	0	0
Day 4+CM	23	62	5	8	NT	NT	2	0	0	0	0
Day 4-CM	28	57	7	5	NT	NT	3	0	0	0	0
Day 8+CM	28	56	7	6	98	96	3	0	0	0	0
Day 8-CM	25	58	8	6	100	97	3	0	0	0	0
Day 10+CM	22	62	8	7	NT	NT	1	0	0	0	0
Day 10-CM	28	57	6	8	NT	NT	1	0	0	0	0

Fresh cells, obtained by leukapheresis, were smeared and stained with Wright-Giemsa. Cultured cells at passage 9 were seeded in Falcon T30 flasks at 2.5×10^5 per ml on day 0 in the presence or absence of CM. On the days after seeding as noted, the cells were mounted by centrifuging 0.5-ml aliquots of the cell suspension at 500 r.p.m. for 5 min in a Shandon-Elliott SCA-0030 cytospin centrifuge. After Wright-Giemsa staining, differential cell counts were performed. Peroxidase and Sudan black stains were performed according to techniques referenced in the text. NT, not tested.

*Segs, segmented neutrophils; †bands, banded neutrophils; †metas, metamyelocytes.

flasks were supplemented with conditioned media (CM) derived from six different monolayer cultures of various first trimester and early second trimester foetal organs as previously described⁶, while control flasks contained no CM. After three weeks incubation leukocyte growth was noted in flasks containing conditioned media from two different human embryonic fibroblast cultures but not in those without CM or containing other sources of CM (Fig. 1). After splitting the leukocytes and further incubation, cell growth continued only in those flasks containing CM from a single lung embryo culture designated DHL (Fig. 1). After several more weeks, however, the cells continued to grow and differentiate along the myeloid series in the absence of any CM supplements. The cells required foetal calf serum for growth, the optimum concentration being 20% in our culture conditions. When seeded in T30 flasks at 2.5×10^5 cells per ml the culture reached a saturation density of $2.5\text{--}3.0 \times 10^6$ cells per ml in 8–9 d with a doubling time of 55–60 h (Fig. 2).

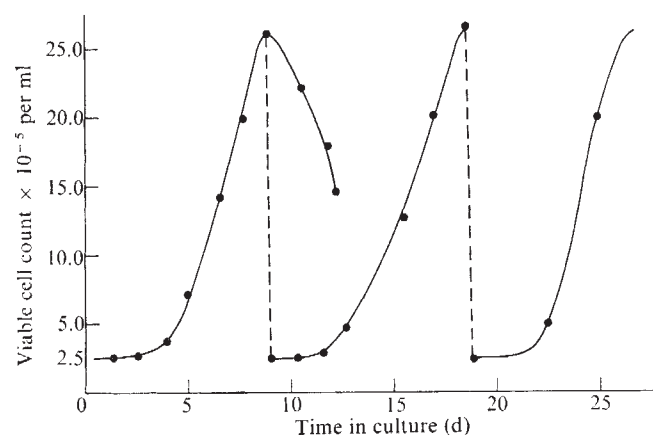
No evidence of mycoplasma contamination as determined by growth assay in agar⁷ has been present. Reverse transcriptase assays of the supernatant culture media⁸ as well as thin-section electron microscopy of fixed-cell pellets showed no evidence of type-C viral production or expression.

Most of the cells when stained by the Wright-Giemsa procedure were myeloblasts and promyelocytes with azurophilic granules⁸, but more mature myeloid cells (myelocytes, metamyelocytes, bands and segmented neutrophils)

were also seen (Fig. 3a,b and Table 1). A small percentage of the cells resembled monocytes with blunt pseudopods, pale staining cytoplasm, and convoluted nuclei (Fig. 3c and Table 1). Differential counts of the cultured leukocytes remained quite similar throughout the growth phase of the cells (Table 1). Histochemical studies using stains specific for myeloid cells confirmed the myeloid nature of the cell culture with over 95% of the cells staining positive for Naphthol ASD chloroacetate esterase⁹, Sudan black¹⁰, and peroxidase¹¹ (Fig. 3d and Table 1). Although the patient's peripheral blood leukocytes from which this culture was derived contained a significant number of lymphocytes (Table 1), no cells resembling lymphocytes are seen in the present culture. In agreement, T- and B-lymphocyte markers as determined by rosette formation with sheep erythrocytes¹², and complement-bound antibody-coated erythrocytes¹³ respectively, were negative. Tests for Epstein-Barr virus nuclear antigen (EBNA)¹⁴ were also negative. Nucleated red blood cells, present in the patient's peripheral blood cells (Table 1), were not seen in the cultured cells and benzidine stains for haemoglobin¹⁵ were negative. Similar morphological and histochemical differentiation has been noted throughout 30 passages of the cell culture over a period of 9 months.

Chromosome studies of the cultured cells (passage 6) revealed aneuploidy with 27 of 40 metaphases examined showing 44 chromosomes and the remaining metaphases varying from 43 to 47 in number. A modal chromosome number of 44 was also observed in short-term cultures of fresh bone marrow obtained from the patient before therapy (J. M. Trujillo, personal communication). These cytogenetic studies together with the morphological similarity between the HL-60 cells and the patient's leukaemic cells strongly suggest that the cells currently being cultured are of leukaemic cell origin. More detailed cytogenetic studies are in progress to confirm this point.

It is unclear exactly what part the conditioned medium played in the origin of the HL-60 cell culture. Attempts to duplicate the experiment using frozen stocks of the patient's fresh uncultured peripheral blood leukocytes were not successful. Also the DHL-conditioned medium failed to stimulate growth of samples of peripheral leukocytes obtained from 18 different myelogenous leukaemia patients, although none of these was of the promyelocytic variant. With the readdition of CM to the growing HL-60 culture, no changes in morphological appearance, histochemical staining, degree of differentiation (Table 1), growth rate, or saturation density were noted. The initial growth of the HL-60 cells only in those flasks containing CM could have

Fig. 2 Continued growth of HL-60 cells in the absence of conditioned medium. The dotted lines indicate a culture split.

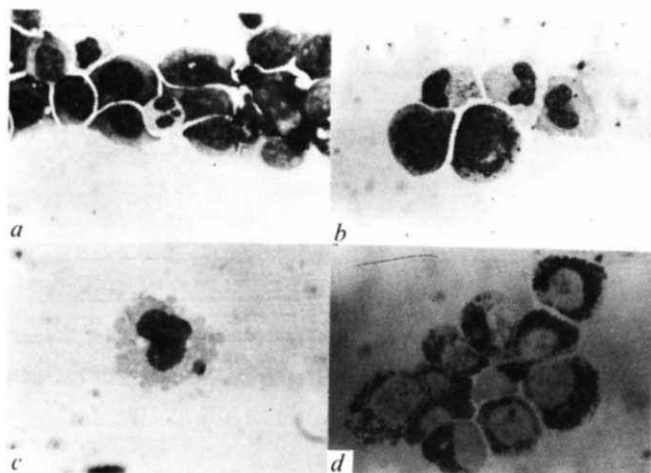


Fig. 3 Cytological and histochemical characteristics of HL-60 cells cultured in the absence of conditioned medium. *a*, Mature polymorphonuclear leukocyte surrounded by myeloblasts and promyelocytes containing cytoplasmic granules. Wright-Giemsa stain ($\times 1,000$). *b*, Two promyelocytes with prominent cytoplasmic granules together with three metamyelocytes undergoing nuclear differentiation. Wright-Giemsa stain ($\times 1,000$). *c*, Single cell with blunt pseudopods, pale staining cytoplasm, and convoluted nucleus resembling a monocyte. Wright-Giemsa stain ($\times 1,000$). *d*, Positive myeloperoxidase stain which histochemically confirms the myeloid morphological appearance of the cells ($\times 1,000$).

been fortuitous, although further studies are required to assess the interaction of the DHL-conditioned medium with the HL-60 cells as well as with leukaemic cells from other patients with promyelocytic leukaemia.

The presence of a cell culture that clearly differentiates along the myeloid series apparently without the addition of an exogenous source of inducer raises some intriguing questions regarding myelopoiesis in normal and leukaemic cells. It is generally assumed that normal stem cells will differentiate along the myeloid series only when influenced by specific microenvironmental factors, which may be similar to granulopoietin or colony-stimulating factor (CSF)¹⁶. Previous reports have described differentiation of certain murine and human myeloid leukaemic cells in both suspension^{6,17} and agar^{18,19}, but only in the presence of appropriate conditioned media or of a 'feeder layer' of peripheral blood leukocytes. Hence the HL-60 myeloid cell culture described here seems to be unique in its ability to proliferate and differentiate without an exogenous source of inducer. These cells may indeed be autonomous with respect to differentiation. It is conceivable, however, that a certain subpopulation of these cultured cells is secreting factors regulating proliferation and differentiation of the remaining cells. Our investigations are directed towards isolating such a subpopulation as well as determining the response of the HL-60 cell culture to the known inducers of myeloid differentiation²⁰.

We thank Dr C. Tai of Biotech Laboratories, Rockville, Maryland who performed the initial cytogenetic studies on the HL-60 cell culture.

STEVEN J. COLLINS
ROBERT C. GALLO
ROBERT E. GALLAGHER

Laboratory of Tumor Cell Biology,
National Cancer Institute,
Bethesda, Maryland 20014

Received 10 June; accepted 26 September 1977.

1. Nilsson, K. & Pontén, J. *Int. J. Cancer* **15**, 321-341 (1975).
2. Lozzio, C. B. & Lozzio, B. B. *Blood* **45**, 321-334 (1975).
3. Klein, E. et al. *Int. J. Cancer* **18**, 421-431 (1976).
4. Lozzio, C. B., Lozzio, B. B., Yang, W.-K., Ichiki, A. T. & Bamberger, E. G. *Cancer Res.* **36**, 4657-4662 (1976).
5. Barak, Y., Shore, N. A., Higgins, G. R. & Vadakan, V. V. *Br. J. Haemat.* **27**, 543-549 (1974).

6. Gallagher, R. E., Salahuddin, S., Hall, W. T., McCredie, K. B. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4137-4141 (1975).
7. Hayflick, L. *Texas Rep. Biol. Med.* **23**, 285-303 (1965).
8. Bernard, J., Lasneret, J., Chome, J., Levy, J. P. & Boiron, M. *J. clin. Path.* **16**, 319-324 (1963).
9. Yam, L. T., Li, C. Y. & Crosby, W. H. *Am. J. clin. Path.* **55**, 283-290 (1971).
10. Williams, W. J. in *Hematology* (eds Williams, W. J. et al.) 1395 (McGraw Hill, New York, 1972).
11. Williams, W. J. in *Hematology* (eds Williams, W. J. et al.) 1393 (McGraw Hill, New York, 1972).
12. Lay, W. H., Mendes, N. F., Bianco, C. & Nussenzweig, V. *Nature* **230**, 531-532 (1971).
13. Jondal, M., Holm, G. & Wigzell, H. *J. exp. Med.* **136**, 207-215 (1972).
14. Reedman, B. M. & Klein, G. *Int. J. Cancer* **11**, 499-520 (1973).
15. LoBue, J., Dornfest, B. S., Gordon, A. S., Hurst, J. & Quastler, H. *Proc. Soc. exp. Biol. Med.* **112**, 1058-1062 (1963).
16. Moore, M. A. S. in *Recent Advances in Cancer Research: Cell Biology, Molecular Biology and Tumor Virology* (ed. Gallo, R. C.) (CRC, Cleveland, Ohio, in the press).
17. Fibach, E. & Sachs, L. *J. Cell Physiol.* **86**, 221-230 (1975).
18. Ichikawa, Y. *J. Cell Physiol.* **74**, 223-234 (1969).
19. Moore, M. A. S., Spitzer, G., Williams, N., Metcalf, D. & Buckley, J. *Blood* **44**, 1-18 (1974).
20. Krystosek, A. & Sachs, L. *Cell* **9**, 675-684 (1976).

Non-random chromosome gains in human lymphoblastoid cell lines

HUMAN lymphoblastoid cell lines derived from the peripheral blood lymphocytes of healthy donors are commonly diploid when examined in the early months after establishment but acquire chromosome abnormalities on prolonged culture. Other lines, notably those derived from Burkitt's lymphoma tissue, may display chromosome aberrations from the outset¹⁻⁵. A partial translocation 8q-14q+ has been demonstrated in the majority of Burkitt lymphoma-derived lines^{5,6} and data on the karyotypes of some human tumours suggest that non-random gains and/or losses of chromosomes may be a feature, in particular, of certain leukaemias and lymphomas⁷⁻¹³. As the emergence of an aneuploid clone from a previously diploid lymphoblastoid line may be associated with other changes suggesting the development of a more 'malignant' phenotype¹⁴, it is relevant to compare the chromosome aberrations detected in such lines with the human tumour data. We have therefore undertaken a study of banded karyotypes of eighty EB virus-carrying human lymphoblastoid lines. The first stage of this analysis is concerned only with gains and losses of whole chromosomes or chromosome arms and the data presented here establish that, considering all the lines together, there have been non-random gains of five autosomes (numbers 3, 7, 8, 9 and 12) and of the sex chromosomes.

Most of the cell lines studied were originally established in this laboratory and all had been maintained in this laboratory for at least 6 months. The derivation of the lines and their 'age' *in vitro* are set out in Table 1. Chromosome analysis was carried out on photographs from quinacrine-stained metaphase spreads¹⁵. (A more detailed examination of chromosome breaks and translocations in these lines is in progress making use of G and R banded, in addition to quinacrine stained, preparations.) In most instances, serial studies have been made on the lines and the emergence of new chromosomally marked clones has been traced over a period of up to 6 yr³. Aberrations detected in only a single cell from a given preparation were ignored while any which recurred in successive clones (identified by additional chromosome markers) from a given line were scored only once for that line. Where a lymphocyte donor had a constitutional chromosome abnormality (Trisomy 21 or XXY, for example) the corresponding abnormality in the derived lymphoblastoid line was not scored as an aberration. There was rarely any difficulty in recognising duplication or deficiency of a whole chromosome or chromosome arm and in establishing that such a marker characterised at least a significant proportion of the cells of a given line. The number of photographs in which a particular aberration could be identified was usually between 6 and 20.

The findings are presented in Fig. 1: 36 lines contribute to these data since eight lines listed as 'aneuploid' in Table 1

Table 1 Lymphoblastoid cell lines studied

Derivation	No. of lines	Age < 1 yr		Age > 1 yr	
		Diploid†	Aneuploid	Diploid	Aneuploid
No lympho-reticular malignancy‡	55	23	2	5	25
Lympho-reticular malignancy§	17	1	0	2	14
Burkitt's lymphoma	8	0	0	0	8
Total	80	24	2	5	49

**In vitro* 'age' at latest karyotype analysis. Most lines have been followed serially since establishment.

†Status at latest karyotype analysis.

‡Includes lines established from cord bloods, healthy adults, patients with infectious mononucleosis, carriers of constitutional chromosome aberrations (see text), and patients with cancer of bladder or gastrointestinal tract.

§Includes lines from patients with acute and chronic leukaemias, myelofibrosis, and Waldenstrom's macroglobulinaemia. Excludes Burkitt's lymphoma.

had internal rearrangements without substantial gains or losses and a further seven (including three from Burkitt's lymphoma) were near-tetraploid.

It is evident that gains occur much more frequently than losses. In fact apparent absence of a chromosome has always been associated with the presence of one or more unidentifiable abnormal chromosome whose total length is at least equal to that of the material otherwise unaccounted for. The only exception to this rule is the Y chromosome which was lost from two lines. The residual stemline in one is 45,X. The other was derived from a 47,XXY Klinefelter

syndrome patient and the two X chromosomes remain. In the same line there were gains of chromosomes 3, 8, 11 and 12.

Analysis of the total data shows that trisomy or partial trisomy for five autosomes and the sex chromosome pair occurs with remarkable frequency. Out of a total of 66 additional chromosomes or chromosome arms, more than two-thirds (47) are accounted for by only 6 of the 23 pairs, numbers 3, 7, 8, 9, 12 and X+Y (treating the sex chromosomes as one pair). On the null hypothesis that all pairs contribute equally to the additional chromosomes or chromosome arms, it can be shown that the probability that six pairs account for at least two-thirds of them is less than 10^{-6} . No single line is trisomic for all of these chromosomes but in four cases trisomy 12 is associated with an extra number 3, 8 or 9 and in one other line (mentioned above) there is trisomy for three of the five autosomes.

The cultures examined include multiple (two to four) lines derived from the blood lymphocytes of each of seven donors. Given that the distribution of extra chromosomes in the total series of cultures is non-random, there seems to be no particular tendency for identical aberrations to recur in multiple lines from a single donor. Eight other lines were grown from the blood lymphocytes of patients with constitutional chromosome aberrations (45,X, 47,XXY, 49,XXXXY, 47,XYY, trisomy 21 and balanced autosomal translocations). As a group they have shown no greater or lesser tendency to develop further aneuploidy *in vitro* than lines from chromosomally normal donors.

Trisomy 7 has been noted in at least two fresh Burkitt lymphoma biopsies and in previously-described Burkitt lymphoma cell lines⁵. Our data support the view that this aberration has a particular association with Burkitt's lymphoma since trisomy 7 (in one case partial) was found in four of the five Burkitt lines which could be scored, Daudi, Namalva, Jijoye and RAJI. Three other lines, EB₁, EB₂ and EB₃ are near tetraploid and rather variable in chromosome complement while EB₁ has four large abnormal chromosomes any of which may include a substantial part of an additional number 7. Our results show, however, that trisomy 7 is not restricted to Burkitt-derived lines and the same aberration has also been reported in two lymphosarcoma biopsies, in lymphoblastoid lines from lymphosarcoma and Hodgkin's disease and in cultures of melanoma tissue⁶⁻⁸.

Trisomy 8 has been observed with striking frequency in the acute phase of chronic myeloid leukaemia, in acute non-lymphatic leukaemia, in polycythaemia vera and in other myeloproliferative disorders. In all of these conditions except myeloid leukaemia, trisomy 9 also occurs more often than would be expected if extra chromosomes were distributed at random⁹⁻¹². There is thus some suggestion of concordance between specific chromosomes gains observed in cultured human lymphoblastoid lines and in malignancy, particularly when it affects the lympho-reticular system. On the other

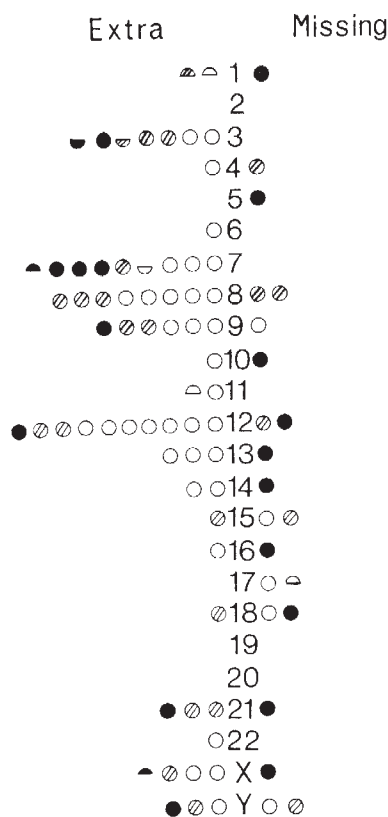


Fig. 1 Cumulative chromosome gains and losses in 36 lymphoblastoid lines. Each line trisomic for a given chromosome is recorded as a full circle to the left of that chromosome number; monosomy is indicated by a full circle to the right. Lines are grouped as in Table 1. Open symbols, lines from patients with no lymphoreticular malignancy; cross-hatched symbols, lines from patients with lymphoreticular malignancy, not Burkitt's lymphoma; filled symbols, Burkitt's lymphoma-derived lines. Partial trisomy or monosomy indicated by upper semicircle (short arm) or lower semicircle (long arm). In addition to gains or losses indicated, there were 19 unidentifiable abnormal chromosomes in the total sample, nine from four Burkitt's lymphoma-derived lines, six among three lines from patients with non-Burkitt's lymphoma lymphoreticular malignancy and one each in four other lines.

hand, trisomy 3 or 12 do not seem to be characteristics of any human neoplasms reported so far and in acute leukaemias loss rather than gain of one chromosome seems to be the norm^{11,12}. Furthermore, there are several examples of non-random chromosome gains or losses in particular human tumours which do not correspond to the changes observed in our material¹⁰⁻¹³.

The phenomenon of non-random chromosome gains in cultured human lymphoblastoid lines is established by this study. Its significance may be tested by planned experiments in xenotransplantation involving diploid and aneuploid cells from the same lines.

We thank Professor M. A. Epstein (Bristol) for provision of cell lines, Mr Norman Davidson for photographic work, Mr Andrew Carothers for statistical analysis and Miss Karin Buckton for useful discussions.

C. M. STEEL
MARILYN A. WOODWARD
CATHERINE DAVIDSON
JUDITH PHILIPSON
ELIZABETH ARTHUR

MRC Clinical and Population Cytogenetics Unit,
Western General Hospital,
Crewe Road, Edinburgh, UK

Received 24 June; accepted 29 August 1977.

1. Nadkarni, J. S. *et al.* *Cancer* 23, 64 (1969).
2. Macek, M. *et al.* *Cancer Res.* 31, 308 (1971).
3. Steel, C. M., McBeath, S. E. & O'Riordan M. L. *J. natn. Cancer Inst.* 47, 1203 (1971).
4. Bloom, A. D., McNeill, J. A. & Nakamura F. T. in *Chromosomes and Cancer* 565 (ed. German, J.) (Wiley, New York, London, Sydney and Toronto, 1974).
5. Zech, L., Haglund U., Nilsson K. & Klein G. *Int. J. Cancer* 17, 47 (1976).
6. Manolov, G. & Manolova, Y. *Nature* 237, 33 (1972).
7. Fleischman, E. W., Hakansson, C. H., Levan, A. & Moller, T. *Hereditas* 70, 243 (1972).
8. McCulloch, P. B., Dent, O. B., Hayes, P. R. & Liao, S. K. *Cancer Res.* 36, 398 (1976).
9. Rowley, J. D. *Proc. natn. Acad. Sci. U.S.A.* 72, 152 (1975).
10. Mitelman, F., Levan G., Nilsson O. G. & Brandt L. *Int. J. Cancer* 18, 24 (1976).
11. Rowley, J. D. *Cancer* 36, 1748 (1975).
12. Mitelman, F. & Levan G. *Hereditas* 82, 167 (1976).
13. Mark, J. *Adv. Cancer Res.* 24, 165 (1977).
14. Nilsson, K. *Haematology Blood Transfusion* 20, 253 (1977).
15. Steel, C. M. *Nature* 233, 555 (1971).

Essential role of surface-bound chemoattractant in leukocyte migration

MANY chemotactic factors, usually proteins or peptides, have been isolated and studied, but little is known about the basic mechanism of leukocyte migration. This movement is termed chemotaxis if its direction is determined by substances in the cells' environment¹. The chemotactic agent is assumed to convey information to the leukocytes by interaction with receptors. The subsequent sequence of events thus triggered in the cells is unknown but metabolic changes such as activation of an esterase have been reported as occurring as the cells move forward (for review see ref. 2). A role for surface-bound chemoattractant in cell locomotion was suggested by the observation that mouse fibroblasts move preferentially from substrates of lesser to those of greater capacity to bind to cells^{3,4}. A close relationship has also been suggested between the mechanism of chemotaxis and phagocytosis⁵. In both cases cells form protrusions of motile membrane which carry receptors for either chemoattractants or opsonins. For example, such membrane protrusions have been shown to bind bacteria and also to attach to opsonised particles, through Fc receptors to IgG-Fc or through C3 receptors to C3 fragments. Once initial contact is established during phagocytosis further membrane segments bind, enabling the cell to creep around the bacterium until it is ingested. We report here that using casein, a chemotactic agent in a standard chemotaxis assay⁶, we have found that binding of casein to a solid substratum is essential for its attracting capacity.

Two sets of microporous cellulose nitrate filters were prepared. To the first set ¹²⁵I-casein (casein, from Merck, Darmstadt)

was dissolved in NaOH and adjusted to pH 7.35 by HCl) was bound in a dose-dependent manner (Fig. 1) by incubating the filters for 10 min at room temperature with casein concentrations ranging from 0.5 to 6 mg ml⁻¹. Unbound casein was removed by placing the filters in a suction device and washing with 25 ml of Hank's balanced salt solution (HBSS), after which no additional ¹²⁵I-casein could be eluted. The second set of filters was treated with HBSS only. To test for surface-bound chemotactic factors, the filters were placed in modified Boyden chambers^{7,8}. The upper compartments were filled with a suspension of 2 × 10⁶ guinea pig peritoneal granulocytes, which were collected 18–22 h after injection of glycogen (0.1%, w/v), and the lower compartments were filled with either HBSS (casein-treated filters) or increasing amounts of casein (HBSS-treated filters). In the latter case, the casein concentrations in the lower compartments were the same as those used to prepare the casein-treated filter. After incubation for 90 min at 37 °C the distance migrated by the granulocytes into the filter was measured.

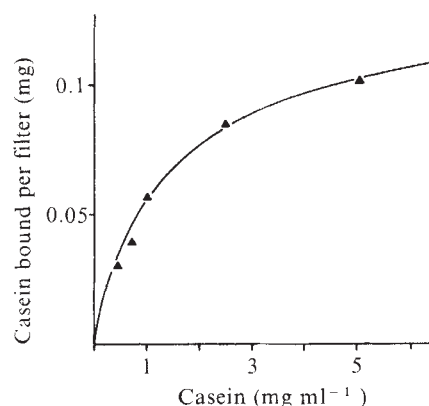


Fig. 1 Uptake of casein by cellulose nitrate filters. Cellulose nitrate filters (3 µm pore size, 25 mm diameter, type AE97, Schleicher & Schüll) were incubated (10 min at room temperature) in 2 ml of solutions containing increasing concentrations of ¹²⁵I-labelled casein. The filters were then placed in a suction device (type FN 023/0, Schleicher & Schüll) and 25 ml of HBSS were passed through the filters. Based on the amount of radioactivity bound per filter, the total amount of casein attached to the filter was determined.

With casein concentrations of approximately 3 mg ml⁻¹, almost identical migration distances were obtained (Fig. 2) whether the filters were preincubated in casein solutions (first set) and used without further addition of casein to the Boyden chamber, or whether the filters were initially free of casein (second set) and used in the test by adding casein to the lower compartment of the chamber. At casein concentrations of 0.5 and 1 mg ml⁻¹ the granulocytes in the first set of filters were found to have migrated slightly longer distances than the granulocytes in the second set of filters. Concentrations of casein higher than 3 mg ml⁻¹ led to opposite results for the two sets of filters. While with the first set of filters a slight increase in migration distance was observed (Fig. 2), the second set yielded a continuous decrease of the migration distance in parallel with the increase in casein concentration in the lower compartment. This reduction distance can therefore only be attributed to the presence of casein which was not bound to the substratum. These experiments were repeated 10 times and showed a clear reproducibility. If purified casein was used for the experiments instead of the commercially available casein, similar results were obtained with filter-bound casein whereas in the standard chemotactic assay higher casein concentrations were required to affect inhibition of migration (data not shown).

To confirm the role of surface-bound chemoattractant, filters soaked with a solution of 2.5 mg ¹²⁵I-casein per ml and washed in HBSS, were incubated for 20 min at 37 °C with solutions of

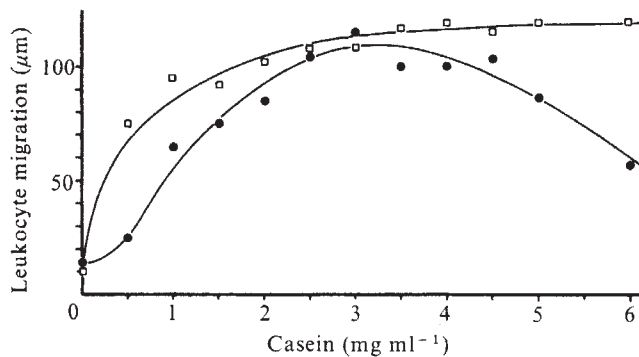


Fig. 2 Effect of filter-bound casein on leukocyte migration. Cellulose nitrate filters prepared as described in the legend to Fig. 1 were used in the Boyden chamber with 1 ml of 2×10^6 guinea pig peritoneal granulocytes in the upper compartment and HBSS in the lower compartment (□). For controls, filters preincubated in HBSS instead of casein were used with solutions of increasing casein concentration in the lower compartment (●). Along the abscissa the casein concentrations are indicated which, in the case of the first set of filters (□), were present during the preincubation and, in the case of the control filters (●), during the actual migration test.

increasing concentrations of papain (1×10^{-8} – 1 mg ml $^{-1}$) and the papain-induced release of casein from the filter was followed (Fig. 3). It was found that increasing concentrations of papain led to increasing release of radioactive material and, simultaneously to decreasing migration distances. In control experiments it was shown that preincubation of the filter with papain, followed by washing with HBSS had no effect on the migration of the leukocytes towards the casein in the lower compartment. Thus we could exclude the possibility that any papain remaining in the filters after the washing procedure would interfere with migration of the cells.

These results suggest that filter bound casein has the essential role in the promotion of casein-induced leukocyte migration, and that casein in free solution may even be inhibitory. Whether this is due to receptor blockade or to toxic effects is not clear. From our experiments with purified casein, the latter possibility seems to be more likely. It is also important to establish whether other chemoattractants behave like casein. If this is a general characteristic, cells responding to a chemotactic stimulus could be envisaged as creeping along by holding on to surface-bound chemoattractants with special receptors or receptive

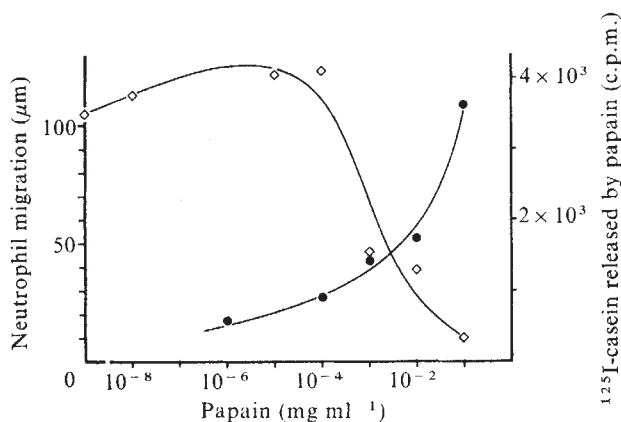


Fig. 3 Reduction of casein-dependent leukocyte migration by pretreatment of the ^{125}I -casein-coated filters with papain. Filters incubated in a solution of 2.5 mg ^{125}I -casein per ml and washed in HBSS were incubated 20 min at 37°C with solutions of increasing concentrations of papain (abscissa) containing appropriate amounts of cysteine. Then the filters were washed and the release of radioactivity was determined (○). The same filters were tested for leukocyte migration (◇).

sites. Binding of fluid phase reagents to the same cell receptors might result in blocking the interaction of these receptors with the surface-bound material. Also, fluid-phase reagents might bind first to the cell receptors, and thereby promote secondary binding to a nearby surface, or they might stimulate cells to secrete factors with functional characteristics similar to those of casein. This is not so far fetched, as we have observed that supernatants obtained by incubating leukocytes for 30 min at 37°C in HBSS contain substances which, like casein, modulate filters so that they now can induce the locomotion.

According to this view, leukocytes migrating *in vivo* would be crawling along tissue surfaces coated with chemoattractants. By analogy with fibroblasts^{3,4} the cells would migrate to where there was a greater amount of attractive substances fixed per unit surface area which would coincide with a tighter binding. This is supported by preliminary experiments with filters to which casein was not attached uniformly but with an increasing concentration from one side to the other (unpublished results). This movement would only be possible if unoccupied receptors could be made available continuously at the leading edge of the cell or on the protrusions actively reaching out from the cell body searching for sites of attachment. Therefore the cells must either be able to synthesise new receptor sites or to re-use the already existing ones. Harris's concept⁹ that the surface proteins in moving cells continuously circulate within the membrane from the front of the cell to its trailing end, then become internalised and, after a hypothetical passage through the cytoplasm, become incorporated again into the front section of the cell, seems to be most attractive. It is possible that proteolytic activity might be involved in freeing the receptors from the bound substance, since chemotactic factors are known to induce release of lysosomal enzymes^{10–12}.

We thank Dr P. Henson for helpful discussion and Miss Jutta Scholtka and Miss Helga Braun for technical assistance. The work was supported by the DFG (SFB 107 Di A5 and Ti 63/5).

MANFRED P. DIERICH
DIETER WILHELM

*Institute for Medical Microbiology,
Johannes Gutenberg-University Mainz,
Augustusplatz, D-65 Mainz*

GERD TILL

*Institute for Immunology and Serology,
Im Neuenheimerfeld 305,
Postfach 103060,
D-69 Heidelberg, FRG*

Received 30 May; accepted 15 August 1977.

- Keller, H. U. *et al. Clin. exp. Immun.* 27, 377–380 (1977).
- Wilkinson, P. C. *Chemotaxis and Inflammation* (Churchill and Livingstone, Edinburgh and London, 1974).
- Carter, S. B. *Nature* 208, 1183–1187 (1965).
- Harris, A. *Expl. Cell Res.* 77, 285–297 (1973).
- Griffin, F. M., Griffin, J. A., Leider, J. E. & Silverstein, S. G. *J. exp. Med.* 142, 1263–1282 (1975).
- Keller, H. U. & Sorkin, E. *Int. Arch. Allergy appl. Immun.* 31, 575 (1967).
- Boyden, S. V. *J. exp. Med.* 115, 453 (1962).
- Zigmond, S. H. & Hirsch, J. G. *J. exp. Med.* 137, 387 (1973).
- Harris, A. K. *Nature* 263, 781–783 (1976).
- Goldstein, I., Hoffstein, S., Gallin, J. & Weissmann, G. *Proc. natn. Acad. Sci. U.S.A.* 70, 2916 (1973).
- Becker, E. L., Showell, H. J., Henson, P. M. & Hsu, I. S. *J. Immun.* 112, 2047–2054 (1974).
- Showell, H. J. *et al. J. exp. Med.* 143, 1154 (1976).

Membrane-bound H-2 and H-Y antigens move independently of each other

ON exposure to foreign antigens, mouse T lymphocytes can be stimulated to mature into cytotoxic cells able to kill target cells carrying these antigens^{1–4}. Killing is restricted to cells bearing, in addition to the foreign antigen, the same H-2K or H-2D antigens as the stimulating cell. Cells bearing different H-2 antigens are not killed. To kill a lymphocyte, two recognition steps are necessary: one involving the foreign antigen and another involving H-2

antigens. This associative recognition has been demonstrated for many antigens and must be considered a general phenomenon. Two opposing views have been proposed to explain the associative recognition⁵. According to the altered-self hypothesis, the antigen links up on the cell membrane with H-2 molecules and the complex of antigen plus H-2 is then recognised by a single receptor on a T lymphocyte. According to the dual recognition hypothesis, each T lymphocyte has two receptors, one for the antigen and the other for H-2, and it is the combination of the two receptors on each cell that is responsible for the associative recognition. The strongest evidence in favour of the first hypothesis is the report of a physical association between antigens and H-2 (see, for example, ref. 6). The link-up of the antigen to H-2, however, might not be related directly to the associative recognition, but rather reflect some other phenomenon in the membrane, in which case there should be examples of associative recognition without a linkage between the antigens. We present evidence for such an association.

Associative recognition has been demonstrated for H-Y, the male-specific antigen. Gordon *et al.*³ showed that C57BL/10 female mice, when primed to H-Y, develop cytotoxic effector cells which are able to lyse male target cells carrying the same H-2 haplotype as the females (H-2^b). According to the altered-self hypothesis, the H-Y and H-2 should form supramolecular complexes that, in the membrane, redistribute together when exposed to either H-Y or H-2 antibody. We tested this assumption as follows. We incubated C57BL/6 (H-2^b) male thymocytes at 37 °C with either anti-H-2.33 or anti-H-2.2 (or both sera) which react with private antigens controlled by the H-2K^b and H-2D^b alleles of the H-2^b haplotype, respectively. We then stained the cells with rabbit anti-mouse IgG (RAMiG) serum conjugated with tetramethylrhodamine isothiocyanate (TRITC)⁷. In these conditions most of the H-2 molecules were in a cap at one pole of the cell. We then applied to these capped cells the mouse anti-H-Y serum at 0 °C and

determined the distribution of the H-Y antibodies using goat anti-mouse IgG (GAMiG) conjugated with fluorescein isothiocyanate (FITC). If H-2 and H-Y were associated, then capping of H-2 should cause capping of H-Y. In this case, no H-Y should remain outside the H-2 caps (the red and green fluorescence of TRITC and FITC, respectively, should overlap and be seen only in the cap with the remainder of the cell not staining at all). If, on the other hand, H-2 and H-Y were not physically associated, there should be a red cap of H-2 and diffuse green staining of H-Y (and—in cases where the GAMiG antibodies find places on the cap unoccupied by the RAMiG—possibly also green cap).

The results are shown in Table 1. When B6 cells were first incubated at a capping temperature (37 °C) with anti-H-2K.33 serum, about one-half stained specifically for the H-2K.33 antigen and in all of them the red TRITC stain had gathered in caps. (The thymocyte population is heterogeneous in terms of the degree of H-2 antigen expression; the weakly H-2 positive cells were classified as negative in the staining conditions used in this experiment.) When these cells were re-exposed to the same antiserum (anti-H-2K.33) in non-capping conditions, no additional staining could be picked up by the FITC-GAMiG antiserum. On exposure to anti-H-2D.2 or to anti-H-Y serum, however, 50 and 60% of the cells, respectively, picked up diffuse green fluorescence. These percentages closely correspond to the percentages of cells reacting with the two antisera in a situation where anti-H-2K.33 was replaced by a medium. We conclude, therefore, that capping of H-2K.33 did not result in any significant capping of H-2D.2 or H-Y. Similar results were obtained when the anti-H-2D.2 serum was used first and the anti-H-2K.33 and anti-H-Y sera second. In this instance the three antigens again redistributed in the cell membrane independently of one another. When the cells were incubated at 37 °C with a mixture of anti-H-2K.33 and anti-H-2D.2 sera, virtually no cells with diffuse green fluorescence were found after re-exposure at

Table 1 Redistribution of H-2 and H-Y antigens after treatment of C57BL/6J male thymocytes with corresponding antisera

First treatment (37 °C)	Second treatment (0° to 4 °C)	No. of cells counted	% Of cells stained with TRITC Redistributed	% Of cells stained with TRITC Diffuse	% Of cells stained with FITC Redistributed	% Of cells stained with FITC Diffuse + caps
Anti-H-2K.33 + TRITC-RAMiG	Anti-H-2K.33	208	100 (48)	0	100 (48)	10 (5)
	+ FITC-GAMiG*					
	Anti-H-2D.2	222	100 (45)	0	24 (11)	111 (50)
	Anti-H-Y	312	100 (32)	0	31 (10)	186 (60)
Anti-H-2D.2 + TRITC-RAMiG	Medium	250	100 (40)	0	100 (40)	15 (6)
	Anti-H-2D.2	200	100 (50)	0	98 (49)	8 (4)
	Anti-H-2K.33	200	104 (52)	0	22 (11)	116 (58)
	Anti-H-Y	230	99 (43)	0	51 (22)	127 (55)
Anti-H-2K.33 + Anti-H-2D.2 + TRITC-RAMiG	Medium	208	100 (48)	0	100 (48)	8 (4)
	Anti-H-2K.33	185	100 (54)	0	100 (54)	11 (6)
	Anti-H-2D.2	160	100 (63)	0	100 (63)	18 (11)
	Anti-H-Y	138	100 (72)	0	41 (30)	94 (68)
Medium	Medium	200	100 (50)	0	106 (53)	8 (4)
	Anti-H-2K.33	100	—	—	9 (9)	77 (77)
	Anti-H-2D.2	100	—	—	12 (12)	79 (79)
	Anti-H-Y	109	—	—	5 (6)	57 (52)
	Medium	100	—	—	2 (2)	0 (0)

Thymuses were removed without parathyroid lymph nodes. Cell suspension was prepared by pressing the organ through a stainless-steel sieve (50-mesh) into Eagle's essential medium (GIBCO) containing 20% heat-inactivated foetal calf serum and cycloheximide (100 µg ml⁻¹), pH 7.4–7.6 (hereafter referred to as 'medium'). Cells were washed twice and resuspended in the medium. For the first treatment, aliquots of thymocyte suspension containing 5 × 10⁶ cells were distributed into small tubes and centrifuged. Alloantiserum or medium, 50 µl, was added to the packed cells, resuspended cells were incubated for another 30 min at 37 °C. After one washing, incubations with alloantiserum (medium) and TRITC-RAMiG were repeated once more. For the second treatment, the cell suspensions were divided into four tubes and transferred to an ice bath. The packed cells were incubated for 30 min with alloantiserum or medium, and, after one washing, with FITC-GAMiG for 30 min. After the last incubation, cells were washed three times at 0° with medium and fixed with 1% paraformaldehyde in phosphate buffered saline. A drop of the cell suspension was placed on microscope slide, allowed to dry, and mounted in 90% glycerol. The slides were scored using a Leitz Ortholux II fluorescence microscope equipped with a Ploem epillumination system. Antisera were as follows. Anti-H-2K.33: (B10.D2 × A)F₁ anti-B10. A(5R); anti-H-2D.2: [C3H/HeJ × B10.D2(R107)]F₁ anti-B10. A(4R); anti-H-Y: C57BL/6J anti-C57BL/6J. The H-2 sera were produced as described elsewhere⁸; anti-H-Y serum was prepared by inoculating C57BL/6 females every 2 weeks with 25 × 10⁶ C57BL/6J male spleen cells i.p., and the antiserum was obtained 10 d after the seventh inoculation. The activity of H-2 and H-Y antisera was determined by cytotoxic test against lymphocytes⁹ and spermatozoa¹⁰, respectively.

*In all 'second treatments' cells were incubated with FITC-GAMiG after initial incubation with alloantiserum or medium.

0 °C to either of the two antisera separately: all the detectable H-2 antigens were capped. But the cells still stained diffusely after exposure to H-Y antiserum. The percentage of the cells with green diffuse fluorescence was once again comparable to that in the aliquots treated first with medium and then with anti-H-Y serum. We conclude, therefore, that complete capping of H-2 antigens has no detectable effect on the distribution of the H-Y antigen: the two types of antigen apparently are not linked in the cell membrane. This conclusion, however, requires at least two qualifications. First, it is possible that while the majority of H-Y molecules are not associated with H-2, a small number are; the latter might not be detected by our method. Second, H-2 and H-Y molecules might be associated before the exposure to antibodies and the antibody might break such bonds. Assuming that neither of these two qualifications applies, one has to conclude either that the H-Y associative recognition is a special case and its mechanism is different from associative recognition of other antigens, or that linkage between the antigen and H-2 is not required for the recognition to occur. Since we find little evidence to support the former possibility, we favour the latter. Our finding does not, however, resolve the dispute between altered-self and dual recognition. It is possible, for instance, that what is needed for recognition to occur via a single receptor is that the H-2 and H-Y (or other antigens) come close to each other in the membrane. With the membrane being in a semifluid state, there is ample opportunity for such an event to occur.

We thank Dr Rolf Zinkernagel for discussion. The work was supported by grants from the USNIH.

ROY GEIB

Department of Microbiology,
University of Texas Southwestern Medical School,
Dallas, Texas 75234

ELLEN H. GOLDBERG

Department of Microbiology,
University of New Mexico School of Medicine,
Albuquerque, New Mexico 87131

JAN KLEIN

Department of Microbiology,
University of Texas Southwestern Medical School,
Dallas, Texas 75235

and
Department of Immunogenetics,
Max Planck Institute for Biology,
Tübingen, FRG

Received 13 June; accepted 12 September 1977.

1. Zinkernagel, R. H. & Doherty, P. C. *Nature* **248**, 701–702 (1974).
2. Bevan, M. J. *J. exp. Med.* **142**, 1349–1364 (1975).
3. Gordon, R. D., Simpson, E. & Samelson, L. E. *J. exp. Med.* **142**, 1108–1120 (1975).
4. Shearer, G. N. *Eur. J. Immun.* **4**, 527–533 (1974).
5. Zinkernagel, R. H. & Doherty, P. C. *Nature* **141**, 1427–1437 (1975).
6. Schrader, J. W., Cunningham, B. A. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5066 (1975).
7. Geib, R., Poulik, M. D., Vitetta, E. S., Kearny, J. F. & Klein, J. *J. Immun.* **117**, 1532–1537 (1976).
8. Klein, J., Klein, D. & Shreffler, D. C. *Transplantation* **10**, 309–320 (1971).
9. Klein, J., Hauptfeld, V. & Hauptfeld, M. *Immunogenetics* **2**, 141–150 (1975).
10. Goldberg, E. H., Boyse, E. A., Bennett, D., Scheid, M. & Carswell, E. A. *Nature* **232**, 478–480 (1971).

Rate of offset of action of slow-acting muscarinic antagonists is fast

ATROPINE generally behaves as a competitive antagonist of acetylcholine (ACh) at muscarinic receptors^{1–9} but its onset and offset of action are slow. It has been suggested that the rates of onset and offset of atropine blockade are determined by the rates of binding to and dissociation from the ACh receptor^{4,5,10} a suggestion which at first sight seems incompatible with competitive blockade. An alternative explanation is that access of antagonist (in contrast to agonist) to the receptors, and escape from their vicinity on washing, is delayed or impeded in some

way^{11–15}. Experiments in which atropine is applied to a small superficial region of the tissue would seem to offer the best chance of reducing this access factor to a minimum and help in deciding which explanation for atropine's slow action is correct. The iontophoretic method presents itself as the method of choice, but until now, no such experiments have been carried out. The results of this type of experiment carried out on smooth muscle are described here. They reveal that the rate of offset of blockade by slow acting muscarinic antagonists can be more than ten times faster than that seen when antagonists are applied by their addition to the solution bathing (or perfusing) the muscle.

The iontophoretic application of ACh (typically using a 50-nA, 2.5-nC pulse) usually elicited a monophasic depolarisation with a latency of 0.1–1 s (Fig. 1) although bi- and polyphasic responses were also observed¹⁸. The sensitivity to ACh was tested every 16–20 s, or occasionally every 8 s, and in stable preparations the response varied little. Three muscarinic blocking agents were studied for their effects on the ACh (or carbachol) response. These were: atropine sulphate, (2-benziloyloxyethyl) diethylmethyl ammonium iodide (the diethyl analogue of lachesine¹⁹, DE lach) and the partial agonist hexyl trimethyl ammonium iodide (hexyl TMA)^{2–14}. DE lach resembles lachesine except that one of the quaternary nitrogen methyl groups is replaced by an ethyl group. This marginally increases its potency¹⁹. The release of any one of these substances from the other barrel of a double-barrelled iontophoretic pipette could reduce or abolish the response to ACh or other muscarinic agonist (Fig. 1). If no further blocking agent was applied, the responses recovered their former size at a rate which was different for each drug.

Table 1 Rate of offset of blockade by muscarinic antagonists

Antagonist	No. of experiments	Time (s)*		$K_{off}(s^{-1})$
		Half recovery	Full recovery	
Atropine	14	62 ± 15	216 ± 24	~ 11 × 10 ⁻³
DE lach	10	21 ± 1.6	46 ± 2.2	~ 34 × 10 ⁻³
Hexyl TMA	1	9	25	~ 77 × 10 ⁻³

*Mean ± s.e.m.

Pulses of antagonist were chosen which produced initially 90–100% reduction in the response to ACh or other agonist. To obtain a quantitative measure of the speed of offset of antagonism, the times from first re-appearance of the response to 50% recovery or to full recovery, were measured in a number of experiments (Table 1). The effects of an atropine or DE lach pulse sufficient initially to nearly or completely abolish the response to ACh were completely absent on average after about 3.5 min and 45 s,

Table 2 Summary of rate constants of offset of antagonist action as measured by various workers

Antagonist	$K_{off}(s^{-1})$	Tissue	Reference
Atropine	0.3 × 10 ⁻³	Atria	12
	1.3 × 10 ⁻³	Perfused heart	12
	1.8 × 10 ⁻³	Longitudinal muscle ileum	5
	0.3 × 10 ⁻³	Longitudinal muscle ileum	4
	0.9 × 10 ⁻³	Longitudinal muscle ileum	13
	11 × 10 ⁻³	Taenia	This work
Lachesine	3.7 × 10 ⁻³	Longitudinal muscle ileum	5
	0.7 × 10 ⁻³	Longitudinal muscle ileum	13
DE lach	34 × 10 ⁻³	Taenia	This work

respectively, and recovery to half size was achieved at about 1 min and about 20 s.

To facilitate comparison with the work of others, the time to half-recovery has been multiplied by the factor 1.44 to obtain the rate constant of offset of antagonism (K_{off}), assuming an exponential decline of occupancy by blocking agent and a linear relationship between response and agonist dose¹². It is doubtful if these assumptions are strictly justified but the error involved is insignificant when compared with the rates of offset of blockade

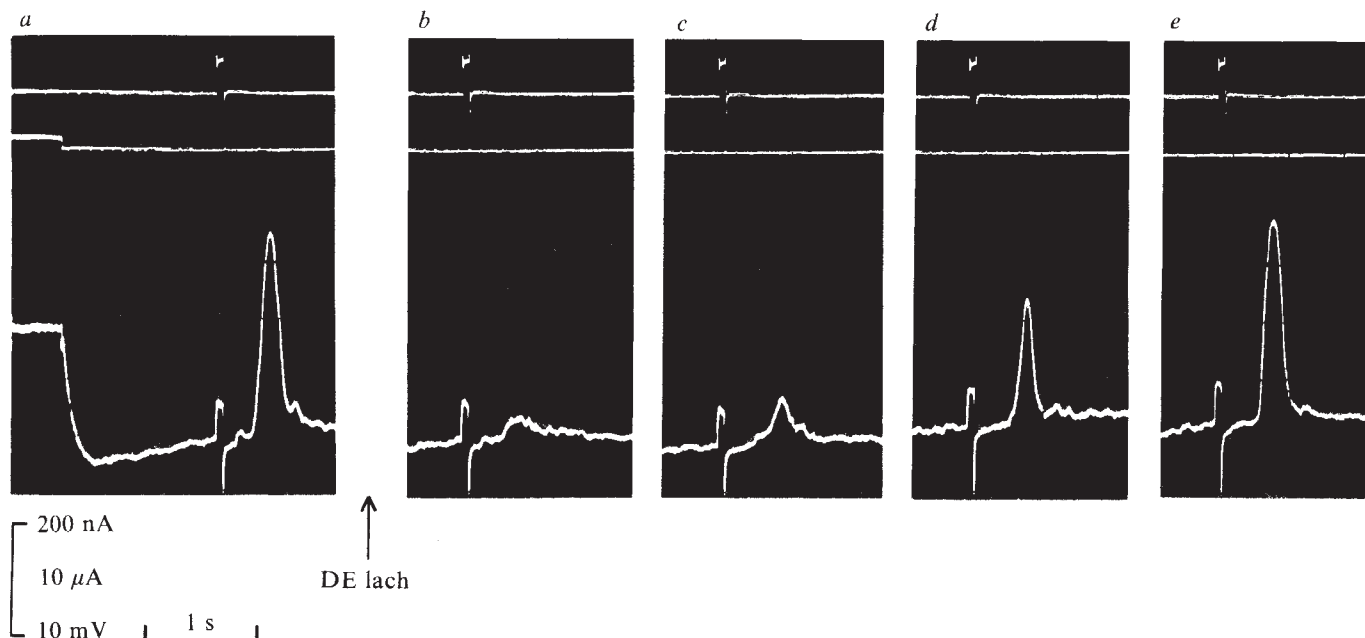


Fig. 1 Block of response to iontophoretic ACh by an iontophoretic pulse of DE lach and its subsequent recovery. *a*, Lowest trace (10 mV calibration). Depolarisation caused by the stimulation of the muscarinic receptors of smooth muscle by ACh (3.6 nC) released as indicated by the top iontophoretic monitor trace (200 nA calibration). Membrane potential was recorded intracellularly by microelectrode within 25 μ m of the point of application of ACh from one barrel of a twin-barrelled iontophoretic pipette with spacer¹⁶. The membrane potential before each application of ACh was first hyperpolarised by passing inward current (middle trace and 10 μ A calibration) to prevent spiking and contraction. This part of the record is shown in *a* but is omitted in subsequent ones. Later responses to ACh were obtained at 8 s, *b*; 24 s, *c*; 40 s, *d*; and 56 s, *e* after DE lach (8 nC, not shown) which was applied from the other barrel of the iontophoretic pipette. The preparation was a narrow longitudinal strip cut from the longitudinal muscle of the guinea pig caecum (taenia). The region of active muscle was further reduced by creating a 'node' of active muscle, perfused with isotonic physiological salt solution (composition given in ref. 17) at 35 $^{\circ}$ C, between two streams of de-ionised sucrose solution in a type of double sucrose-gap apparatus described previously¹⁷.

given in the literature (Table 2), using muscle preparations bathed or perfused with solution containing the blocking agent. In general, previous measurements of the rate of offset of blockade are at least ten times slower (range 6–49) which argues strongly that in experiments described by others it is factors other than just dissociation from the receptors, which determines the rate of offset of blockade.

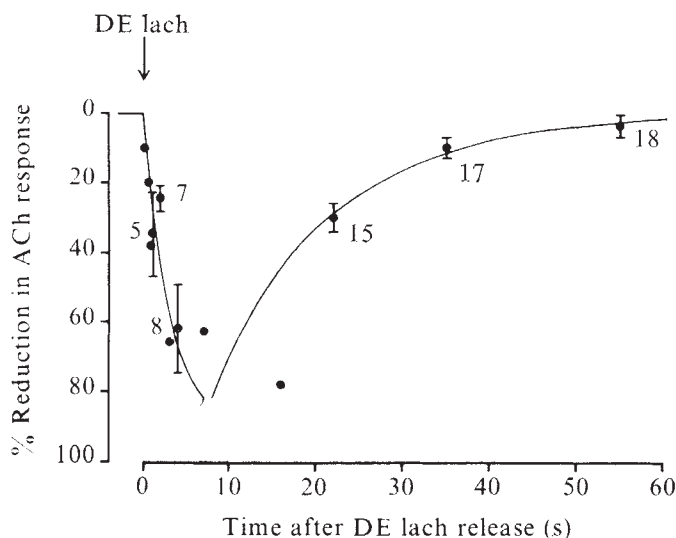


Fig. 2 Time course of onset and offset of blockade by DE lach pulse (16 nC) constructed by combining measurements of the percentage reduction in ACh responses by some 39 separate DE lach pulses placed with varying time relationships to the ACh pulses. Standard errors ($\pm$) are shown only for those points which are the mean of measurements made on at least five responses. The lines (fitted by eye) are exponentials ($\tau = 3.7$ and 13 s). Notice that maximum block (about 85%) by DE lach occurs some 5–10 s after release of DE lach from the pipette.

The rate of offset of blockade by hexyl TMA was estimated by means of a series of applications of hexyl TMA placed in varying time relationships to agonist responses elicited every 20 s, and the percentage reductions in agonist response measured. Combination of the results obtained with 17 separate applications of hexyl TMA in one experiment enabled it to be estimated that full recovery took place within about 25 s, and half-recovery within 9 s of application of a blocking pulse of hexyl TMA. The blocking action of hexyl TMA could only be demonstrated with large pulses (50–100 nC). Smaller pulses (10–20 nC) were observed to have only agonist action.

The simultaneous release of blocking doses of hexyl TMA and a strong agonist abolished the response to agonist immediately. This was not the case with DE lach. Figure 2 shows the time course of onset of DE lach blockade calculated from the results of 39 separate DE lach applications to one preparation where the timing of blocking pulses of DE lach was varied with respect to regular ACh pulses. Blockade by DE lach took some 5–10 s to become maximal and then offset of blockade commenced. The onset of blockade by atropine could not be studied in detail because it was generally found necessary to pass iontophoretic current for several seconds to release sufficient atropine to produce blockade.

The much greater rate of offset of antagonism by atropine and DE lach seen in the present experiments suggests that the muscarinic ACh receptors involved in the depolarising response may be relatively superficially placed, so that escape of antagonist from the immediate vicinity of the receptors does not determine the rate of offset of blockade. In such a case, the rates of offset of antagonism (K_{off}) seen here give lower limits for the dissociation rate constants (k_2) for the binding of atropine, DE lach, and hexyl TMA of 11×10^{-3} , 34×10^{-3} , and $77 \times 10^{-3} \text{ s}^{-1}$ (Table 1). If it is assumed that the equilibrium constants (k_2/k_1) for atropine and DE lach are 1.1×10^{-9} and $1.4 \times 10^{-9} \text{ M}$ (refs. 4, 5 and 13) then the association rate constants (k_1) will be 10^7 and $2.4 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ respectively. These association rate constants are approaching the theoretical maximum for enzyme-substrate reactions of about $10^8 \text{ M}^{-1} \text{ s}^{-1}$ (ref. 21). This supports the hypothesis that the rate of offset of antagonism observed in these experiments may be

giving for the first time a true indication of the rate of dissociation of antagonist from the muscarinic receptor.

This work was supported by the MRC and owes much to discussions with Professor H. P. Rang and Dr D. Colquhoun.

T. B. BOLTON

St George's Hospital Medical School,
Department of Pharmacology,
Cranmer Terrace, London SW17, UK

Received 15 June; accepted 12 September 1977.

- Clark, A. J. *J. Physiol., Lond.* **61**, 547-556 (1926).
- Stephenson, R. P. *Br. J. Pharmac.* **11**, 379-393 (1956).
- Arunkrishnan & Schild, H. O. *Br. J. Pharmac.* **14**, 48-58 (1959).
- Paton, W. D. M. *Proc. R. Soc. Lond.* **B154**, 21-69 (1961).
- Paton, W. D. M. & Rang, H. P. *Proc. R. Soc. Lond.* **B163**, 1-44 (1965).
- Waud, D. R. *Pharmac. Rev.* **20**, 49-88 (1968).
- Pauling, P. J. & Petcher, T. J. *Nature* **228**, 673-674 (1970).
- Rang, H. P. *Nature* **231**, 91-96 (1971).
- Baker, R. W., Chothia, C. H., Pauling, P. & Petcher, T. J. *Nature* **230**, 439-445 (1971).
- Rang, H. P. *Proc. R. Soc. Lond.* **B164**, 488-510 (1966).
- Furchgott, R. F. *Pharmac. Rev.* **7**, 183-265 (1955).
- Thron, C. D. & Waud, D. R. *J. Pharmac. exp. Ther.* **160**, 91-105 (1968).
- Roberts, F. & Stephenson, R. P. *Br. J. Pharmac.* **58**, 57-70 (1976).
- Ginsborg, B. L. & Stephenson, R. P. *Br. J. Pharmac.* **51**, 287-300 (1974).
- Colquhoun, D., Henderson, D. & Ritchie, J. M. *J. Physiol., Lond.* **227**, 95-126 (1972).
- Del Castillo, J. & Katz, B. *Proc. R. Soc. Lond.* **B146**, 339-356 (1957).
- Bolton, T. B. *J. Physiol., Lond.* **250**, 175-202 (1975).
- Bolton, T. B. *Proc. R. Soc. Lond.* **B194**, 99-119 (1976).
- Barlow, R. B., Scott, K. A. & Stephenson, R. P. *Br. J. Pharmac.* **21**, 509-522 (1963).
- Waud, D. R. *J. Pharmac. exp. Ther.* **158**, 99-114 (1967).
- Gutfreund, H. *Enzymes: Physical Principles* (Wiley, London 1972).

Glutamate as transmitter of hippocampal perforant path

CERTAIN experimental manifestations of plasticity, such as habituation¹⁻³ and long-term potentiation^{2,4-6}, can be readily demonstrated at the perforant path-granule cell synapse in the fascia dentata, and these adaptations are believed to have behavioural counterparts. The perforant path-granule cell synapse may thus serve as a model for investigating the simplest components of behaviour and intellectual function. Detailed studies would, however, require knowledge of the transmitter used by the perforant path fibres. Quite apart from the problem of synaptic plasticity, identification of the perforant path transmitter would add substantially to our knowledge of excitatory transmitters in the central nervous system. Here we present pre- and postsynaptic evidence that these fibres use glutamate as their transmitter.

We have previously shown that endogenous glutamate is released by dentate slices⁷ and hippocampal synaptosomes⁸ in a manner characteristic of transmitter substances and that this release originates in part from the perforant path fibres⁹. These fibres do not release aspartate⁹ or γ -aminobutyrate (GABA)¹⁰. There was also a suggestion⁹, since confirmed¹¹, that the perforant path boutons transport glutamate by a high affinity process. While these observations show that glutamate may be the transmitter at the perforant path-granule cell synapse, one could object that the slices and synaptosomes were depolarised only by chemical, and therefore non-physiological, stimuli and that no evidence has been presented implicating glutamate in the postsynaptic response. We have now shown that electrical field stimulation releases endogenous glutamate from slices of the fascia dentata and that a glutamate antagonist reduces the efficacy of perforant path stimulation.

In eight experiments a longitudinal slice of the internal leaf of the rabbit fascia dentata was electrically stimulated while being continuously superfused. When the slice was superfused with a physiological medium containing 2 mM free Ca^{2+} , electrical stimulation elevated the rate of glutamate release by a maximum of 300-1,200 pmol per 30 s per mg protein. In three experiments one slice was superfused with normal medium and another with Ca^{2+} -free medium supplemented with Mg^{2+} and Mn^{2+} to avoid membrane destabilisation (Fig. 1). Superfusion with a Ca^{2+} -free medium reduced glutamate output in response to stimulation by 62, 78 and 90%. Thus electrical stimulation, like elevated K^{+} and veratridine²,

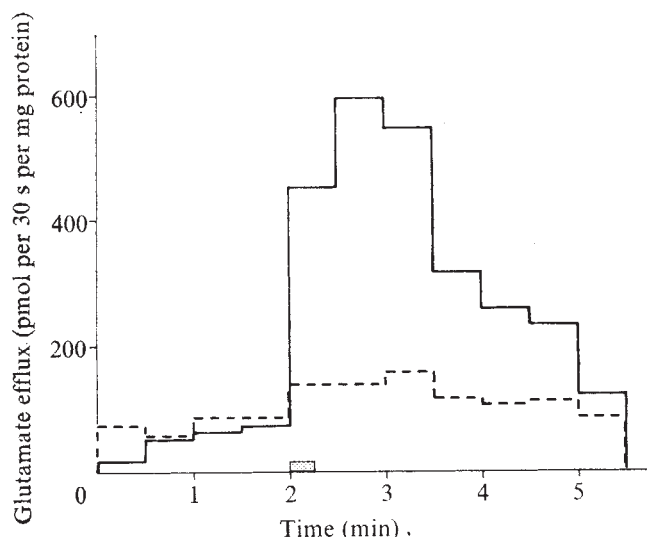


Fig. 1 Time course of glutamate efflux from a slice of fascia dentata in the presence and absence of Ca^{2+} . In each experiment a single 0.35-mm thick longitudinal slice of the free edge (internal leaf) of the fascia dentata was cut from a dissected rabbit hippocampus and immersed in a physiological medium at 37 °C for 5-10 min. Then the slice was superfused with the same 37 °C-medium at 1 ml min⁻¹ in a 0.5-ml chamber equipped with spiral silver electrodes. During the period indicated by the dotted bar the tissue was stimulated at 100 Hz with 10 V square wave pulses of alternating polarity and 5 ms duration. Superfusate fractions were collected continuously, and their glutamate content was determined by a micro-enzymatic fluorometric method¹⁸. The figure shows results from one set of paired experiments. Solid lines, glutamate efflux into normal medium (2 mM free Ca^{2+}); dashed lines, glutamate efflux into Ca^{2+} free medium (with 25 mM MgCl_2 and 0.1 mM MnCl_2).

releases glutamate from the fascia dentata in a Ca^{2+} -dependent manner. Since extracellular Ca^{2+} is required for transmitter release¹², but not for carrier-mediated efflux¹³, this finding agrees with previous evidence⁸ that the glutamate released from fascia dentata by depolarising stimuli originates mainly from presynaptic transmitter stores.

The question remains as to whether the glutamate presumably released by stimulation of the perforant path fibres produces the excitatory postsynaptic potential (e.p.p.) recorded in the granule cell. Proof of this hypothesis requires the demonstration that glutamate and the perforant path transmitter exert physiologically and pharmacologically identical actions on the granule cell. As a first approach to this problem, we have examined the effect of the glutamate analogue, DL-2-amino-4-phosphonobutyric acid (APB), on perforant path responses in the hippocampal slice preparation^{6,14}. APB competitively

Table 1 Effect of APB on averaged responses to perforant path stimulation

	Amplitude of extracellular negativity (% of control)							
	APB				Recovery			
	1	3	5	10 min	1	3	5	10 min
Drug (7)	61 ± 5	51 ± 3	48 ± 4	42 ± 4	56 ± 6	68 ± 6	68 ± 8	74 ± 9
No drug (5)	97 ± 2	94 ± 3	89 ± 4	82 ± 6	82 ± 7	77 ± 7	79 ± 9	76 ± 11

Responses of granule cell populations to perforant path stimulation were recorded as shown in Fig. 2. Amplitudes of these extracellularly recorded potentials were averaged over a period of 50 s at the times indicated. Values are expressed as mean percentages of the averaged potentials recorded within the 50-s period immediately preceding the introduction of APB into the medium ($\pm$ s.e.m. for the number of slices in parentheses). In experiments without APB response amplitudes were averaged at comparable times after the beginning of stimulation.

and reversibly antagonises the action of glutamate iontophoretically applied to invertebrate muscle fibres¹⁵ and in millimolar concentrations inhibits neuromuscular transmission¹⁶. Since invertebrate motoneurons almost certainly use glutamate as their transmitter¹⁷, APB seems to inhibit synaptic transmission mediated by glutamate. Accordingly, we stimulated the perforant path fibres in the presence and absence of 2.5 mM APB and recorded the extracellular field potentials in the molecular layer of the fascia dentata.

The results of one such experiment are shown in Fig. 2. When the perforant path fibres were continuously stimulated at 0.2 Hz, a sizeable negative potential was recorded whose amplitude declined to some extent over a 22-min period (Fig. 2a), probably as a result of habituation^{2,3}. Introduction of APB into the superfusion medium substantially reduced the amplitude of the potential within 1 min (Fig. 2b). The amplitude recovered just as rapidly upon return to normal medium and finally reached the habituated level.

Table 1 summarises the variation of averaged extracellularly recorded potential amplitudes with time in the presence or absence of APB. By comparison with responses in normal medium averaged at similar times after the onset of stimulation, APB was seen to have been maximally effective within 3 min (46% reduction), and 3 min of washout was sufficient to terminate

the effect of the drug. This result indicates, as expected, a readily reversible antagonism.

To determine whether APB exerted merely a local anaesthetic action, we determined its effect on the antidromic potential elicited by stimulation of the mossy fibres (granule cell axons). In four experiments APB did not alter the amplitude or form of the extracellularly recorded potentials, thus ruling out the possibility that APB acted by reducing the excitability of the granule cells. Also in several experiments APB was seen not to affect presynaptic fibre potentials.

APB probably did not act on the presynaptic boutons, since at 2.5 mM it did not reduce the quantity of glutamate released by elevated K⁺ or affect the high affinity transport of glutamate. It, therefore, seems to exert its effect at the postsynaptic level, suggesting a pharmacological identity between glutamate and the perforant path transmitter. This conclusion must be tentative, however, until it can be rigorously shown that APB acts as a glutamate antagonist at this particular synapse.

The results of this and previous studies suggest that perforant path boutons release glutamate from a transmitter store and that the glutamate so released might then depolarise the granule cell. Thus, glutamate is most likely the transmitter used at the perforant path-granule cell synapse. Our results with APB, moreover, suggest that this antagonist will prove a useful tool in dissecting the mechanisms which underlie synaptic plasticity at this site.

We thank Mrs G. Chiang for technical assistance and Miss J. Mueller for typing the manuscript. This work supported by grants from the USNIH and NSF.

W. FROST WHITE
J. VICTOR NADLER
ANDERS HAMBERGER*
CARL W. COTMAN

Department of Psychobiology,
University of California,
Irvine, California 92717

JOSEPH T. CUMMINS

Addiction Research Laboratory,
Veterans Administration Hospital,
Sepulveda, California 91343

Received 27 June; accepted 26 September 1977.

*Permanent address: Institute of Neurobiology, University of Göteborg, Göteborg, Sweden.

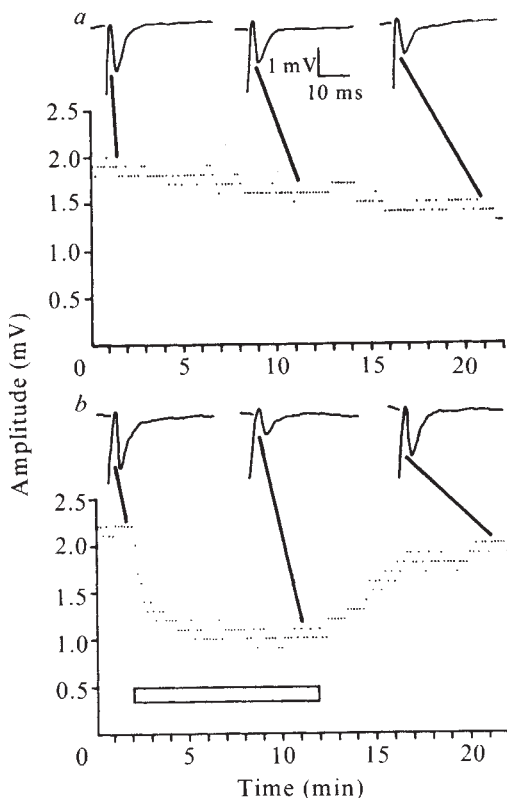


Fig. 2 Reduction of perforant path responses by APB. Slices of 0.7-mm thickness were cut transverse to the long axis of dissected hippocampi, and each was placed on a fritted polypropylene disk within a superfusion chamber. Electrodes were lowered into the chamber, and then the slice was superfused at a rate of 2.5 ml min⁻¹ with Elliott's medium¹⁹ warmed to 32 °C. When the slices showed synaptic activity (1–2 h later), a bipolar stimulating electrode was inserted into the perforant path where it passes through the subiculum and a micropipette recording electrode (3–12 MΩ) into the perforant path terminal zone. The perforant path fibres were stimulated continuously with square wave pulses of 0.05 ms duration at 0.2 Hz and an intensity just below that which evoked a population spike. *a*, Amplitude of every second extracellularly recorded synaptic potential (dots) in a single experiment without drug; *b*, results of an experiment in which 2.5 mM APB was present in the medium during the period denoted by the open bar. Examples of recorded potentials are shown at the times they were evoked.

1. Mays, L. E. & Best, P. J. *Expl Neurol.* 47, 268–279 (1975).
2. Alger, B. E. & Teyler, T. J. *Brain Res.* 110, 463–480 (1976).
3. Teyler, T. J. & Alger, B. E. *Brain Res.* 115, 413–425 (1976).
4. Bliss, T. V. P. & Gardner-Medwin, A. R. *J. Physiol., Lond.* 232, 357–374 (1973).
5. Bliss, T. V. P. & Lomo, T. J. *Physiol., Lond.* 232, 331–356 (1973).
6. Douglas, R. M. & Goddard, G. V. *Brain Res.* 86, 205–215 (1975).
7. Nadler, J. V., White, W. F., Vaca, K. W., Redburn, D. A. & Cotman, C. W. *J. Neurochem.* 29, 279–290 (1977).
8. Sandoval, M. E., Horch, P. & Cotman, C. W. *Brain Res.* (in the press).
9. Nadler, J. V., Vaca, K. W., White, W. F., Lynch, G. S. & Cotman, C. W. *Nature* 260, 538–540 (1976).
10. Nadler, J. V., White, W. F., Vaca, K. W. & Cotman, C. W. *Brain Res.* 131, 141–158 (1977).
11. Storm-Mathisen, J. *Brain Res.* 120, 379–386 (1977).
12. Rubin, R. P. *Calcium and the Secretory Process* (Plenum, New York, 1974).
13. Bennett, J. P., Logan, W. J. & Snyder, S. H. *J. Neurochem.* 21, 1533–1550 (1973).
14. Richards, C. D. & White, A. E. *J. Physiol., Lond.* 252, 241–257 (1975).
15. Cull-Candy, S. G., Donnellan, J. F., James, R. W. & Lunt, G. G. *Nature* 262, 408–409 (1976).
16. Clements, A. N. & May, T. E. *J. exp. Biol.* 61, 421–442 (1974).
17. Gerschenfeld, H. M. *Physiol. Rev.* 53, 1–119 (1973).
18. Graham, L. T. & Aprison, M. H. *Analyt. Biochem.* 15, 487–497 (1966).
19. Elliott, K. A. C. in *Handbook of Neurochemistry* (ed. Lajtha, A.) 2, 103–114 (1969).

Regulation of prolactin release by endogenous opiates

ENDORPHINS, the endogenous peptides recently isolated and identified in brain have been implicated in regulation of pain^{1–4}. But their wide distribution throughout the brain^{5–7} and their profound behavioural effects after central administration⁸ suggest an involvement in other central nervous system processes. Immunohistochemical identification of

these endorphins (refs 9, 10 and F. Bloom, personal communication) in hypothalamic neurones indicated that neuroendocrine effects are probable. Morphine was previously reported to block ovulation, while more recently, morphine¹¹ and endorphins¹²⁻¹⁴ were observed to stimulate release of prolactin and growth hormone. Yet, these observations have not established a direct, tonic participation of endorphins in hypothalamic and anterior pituitary function. The absence of any direct action of opiate antagonists has been taken as an argument against any tonic role of endorphins. Recent reports indicate, however, that opiate antagonists do modify on-going central processes. For example, naloxone and naltrexone lower pain threshold in appropriate conditions in man and experimental animals¹⁵⁻¹⁸. Using naltrexone as a tool to block opiate receptor function, we have explored whether endorphins are tonically involved as a putative neurotransmitter in the regulation of prolactin release. The results presented here demonstrate a new instance where an opiate antagonist modifies normal function. The data agree with a preliminary report indicating that naloxone reduced prolactin release in immature female rats¹⁹.

As shown in Table 1, naltrexone at a dose of 1 mg per kg body weight produced a 50% decrease of serum prolactin concentration of male rats in 20 min. This decrease lasted for more than 2 h and was observed when blood was collected by decapitation or by cardiac puncture. The prolactin decrease elicited by naltrexone was the same in rats with a deafferented hypothalamus (Table 1), so we infer that the inhibition of prolactin release reflects a direct effect of naltrexone at the hypothalamic-pituitary level. Large amounts of β -endorphin are reported to be present in the pituitary²⁰⁻²². We have examined whether naltrexone inhibits prolactin release from isolated pituitary gland *in vitro*. When pituitary halves were incubated without or

with naltrexone in concentrations up to 10^{-4} M, the prolactin release was similar in the treated and untreated halves. The ability of naltrexone to reduce prolactin release stimulated by various mechanisms was also studied. The models used were: (1) chronically elevated prolactin release in oestrogen-primed ovariectomised rats; (2) prolactin release elicited by foot shock in male rats; (3) prolactin release elicited by intrahypothalamic injections of β -endorphin in intact rats or by systemic injection of morphine in deafferented rats; and (4), prolactin release elicited by a dopamine receptor blocker or monoamine depletor.

As shown in Table 1, naltrexone had different effects in these four models. In oestrogen-primed ovariectomised rats, naltrexone at 0.25 mg per kg intraperitoneally (i.p.) produced a significant (48%) inhibition of serum prolactin concentration. Doses of naltrexone of 1 or 5 mg per kg i.p. produced approximately a 60% decrease of hormone secretion. In addition, naltrexone is a very powerful inhibitor of the prolactin release occurring after foot shock (Table 1). It has been reported that intraventricular or systemic injection of β -endorphin or morphine stimulates the release of prolactin in rat^{12,23}. To demonstrate that naltrexone acts in the hypothalamus to interfere with the opiate receptor agonist, we injected β -endorphin into the mediobasal hypothalamus (cannula placement was verified histologically) and naltrexone i.p. The release of prolactin induced by 1 μ g of β -endorphin can be blocked by 1 mg naltrexone per kg, i.p., injected 10 min previously (Fig. 1a). Similarly, the increase of serum prolactin induced by morphine in animals with a deafferented hypothalamus was blocked by naltrexone (Fig. 1b). In contrast, naltrexone failed to inhibit the prolactin release (see Table 1) elicited by haloperidol, a powerful inhibitor of dopamine receptors or by reserpine, a depletor of central monoamine stores.

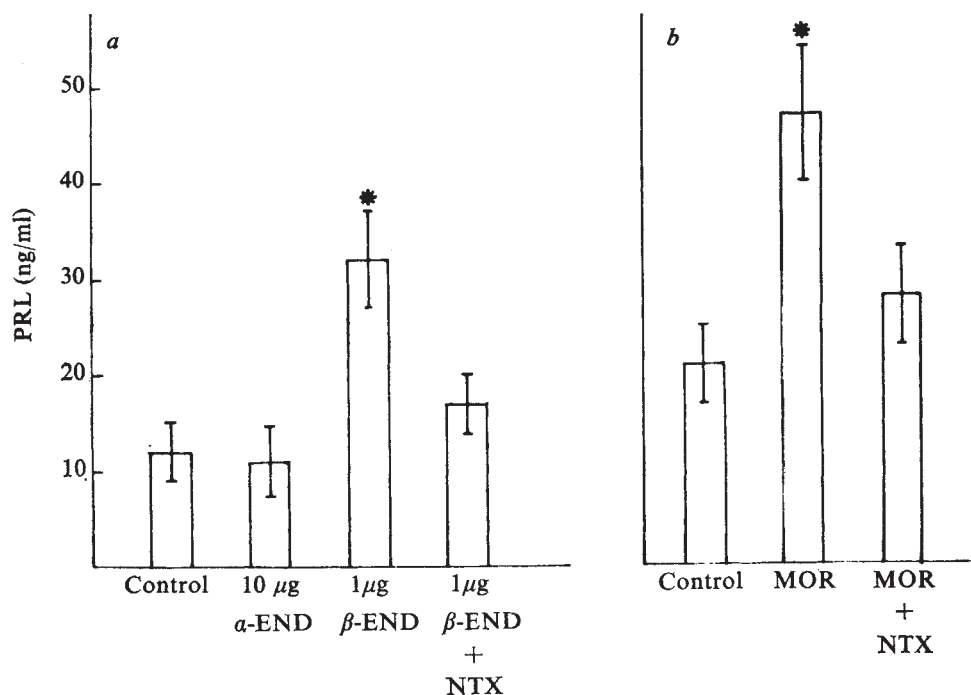
These observations show that β -endorphin stimulates

Table 1 Effects of naltrexone on prolactin (PRL) release

Model	Sex	Method of blood sampling	Naltrexone (mg per kg)	PRL (ng ml ⁻¹)		% Change
				Before treatment	After treatment	
Intact (<i>n</i> = 8)	M	Decapitation	0		11 + 3.2	
			1		3.6 + 1.5*	32
Intact (<i>n</i> = 9)	M	Ether-cardiac puncture	0	28 + 5.4	35 + 12.3	125
			0.25	36 + 5.2	26 + 5.3	87
			1.0	27 + 3.7	18 + 4.4*	54
			5.0	35 + 5.3	17 + 3.8*	49
Hypothalamic deafferentation (<i>n</i> = 6)	M	Ether-cardiac puncture	0	22 + 4.7	19 + 5	91
			1	20 + 3.1	8.7 + 3*	42
Oestrogen primed ovariectomised (<i>n</i> = 6)	F	Ether-cardiac puncture	0	180 ± 66	230 ± 40	126
			0.25	180 ± 49	85 + 53*	52
			1.0	240 ± 44	130 ± 47*	43
			5.0	220 ± 58	90 ± 21*	33
Foot shock (<i>n</i> = 9)	M	Decapitation	0		73 ± 5.4	
			1		31 ± 5.7*	42
Haloperidol (0.25 mg per kg i.p.) (<i>n</i> = 7)	M	Decapitation	0		68 ± 7.9	
			1		55 ± 4.3	81
Reserpine 2 mg per kg i.p. (<i>n</i> = 9)	M	Decapitation	0		44 ± 3	
			1		45 ± 6	102

Intact, Sprague-Dawley, male rats (180–250 g body weight), deafferented male rats 6–15 d after hypothalamic isolation²⁵ or ovariectomised female rats (30 d after surgery) primed with 5 μ g oestradiol benzoate for 5 d were used. Male rats given haloperidol (0.25 mg per kg, i.p.) or reserpine (2 mg per kg, i.p.) were used 45 min and 6 h later. Foot shock (1 mA, 1-s shock per 5 s) was administered for 10 min immediately before killing. Blood samples were collected by decapitation 20 min after i.p. naltrexone or 0.85% NaCl injection or in other cases, by cardiac puncture under light ether anaesthesia before and 20 min after naltrexone or vehicle injection. Serum was separated after overnight clotting at 4 °C. Prolactin was measured using the NIAMD rat prolactin radioimmunoassay kit provided by NIAMD-NIH Pituitary Hormone Program. Mean serum prolactin (PRL) concentration of naltrexone treated rats was compared to control rats and considered significantly different if $P < 0.05$. Treatment effects are described as x serum prolactin concentration of naltrexone treated rats/ x serum prolactin concentration of controls $\times 100$ when rats were decapitated. In experiments where pre- and post-treatment samples were collected, a post-treatment/pre-treatment % was calculated for each rat and mean percentages (post/pre-treatment $\times 100$) were then calculated for each group. In this group statistical significance was evaluated according to the Student's paired *t* test. *n*, No. of animals per group. * $P < 0.05$.

Fig. 1 *a*, Male rats (150–180 g) were implanted with a guide cannula in the mediobasal hypothalamus (4 mm A–P, 0.5 L, 8.6 D–V; König and Klippel²⁶) and injected 3–5 d later with endorphin or vehicle in a volume of 1 μ l in 2 min. Naltrexone (1 mg per kg, i.p.) or 0.85% NaCl was given 10 min before endorphin injection. Blood was collected by decapitation. Cannula position was verified histologically after decapitation. *b*, Male rats (250–300 g) were used 8–16 d after hypothalamic isolation²⁵. Rats were given, systemically, naltrexone in 0.85% NaCl and morphine sulphate or 0.85% NaCl at 10 min and blood was collected by cardiac puncture under light ether anaesthesia. Completeness of deafferentation was determined from histological sections of the brain. PRL, prolactin; α -END, α -endorphin; β -END, β -endorphin; NTX, naltrexone; MOR, morphine.



prolactin release by activation of opiate receptors. More importantly, these experiments indicate that the release of prolactin in physiological conditions can be reduced by a blockade of opiate receptors and suggest that endogenous opiate receptor ligands participate tonically in the regulation of anterior pituitary hormone secretion during basal conditions or when release is stimulated by oestrogen or by excitation of the hypothalamus as during stress.

In agreement with histoimmunological localisation of endorphins (refs 9, 10, and F. Bloom, personal communication) the action of these peptides can be located in the mediobasal hypothalamus. Naltrexone fails to inhibit the prolactin release from the pituitary but it is effective when the mediobasal hypothalamus is attached to the pituitary. It seems that the arcuate nucleus median eminence dopaminergic tract may be involved in the action of opiates since its inactivation pharmacologically prevents the action of naltrexone on prolactin release. Indeed, Ferland *et al.*²⁴ have shown that enkephalin may affect these dopaminergic neurones. It is possible that endorphins and dopamine regulate prolactin release according to an arrangement in which the two neurones act in series. Although further studies are needed to establish the precise role of hypothalamic endorphins in the regulation of prolactin release, these experiments provide strong evidence suggesting their participation.

L. GRANDISON
A. GUIDOTTI

Laboratory of Preclinical Pharmacology,
National Institute of Mental Health,
Saint Elizabeths Hospital,
Washington, DC 20032

Received 16 August; accepted 26 September 1977.

1. Loh, H. H., Tseng, L. F., Wei, E. & Li, C. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2895–2898 (1976).
2. Hughes, J. *et al.* *Nature* **258**, 577–579 (1975).
3. Ling, N. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3308–3310 (1976).
4. Belluzzi, J. D. *et al.* *Nature* **260**, 265–266 (1976).
5. Simantov, R. & Snyder, S. H. in *Opiates and Endogenous Opioid Peptides* (ed. Kosterlitz, H.) 41–48 (Elsevier, Amsterdam, 1976).
6. Wahlstrom, A., Johansson, L. & Terenius, L. in *Opiates and Endogenous Opioid Peptides* (ed. Kosterlitz, H.) 49–56 (Elsevier, Amsterdam, 1976).
7. Yang, H.-Y., Hong, J. S. & Costa, E. *Neuropharmacology* **16**, 303–307 (1977).
8. Bloom, F., Segal, D., Ling, N. & Guillemin, R. *Science* **194**, 630–632 (1976).
9. Elde, R., Hokfelt, T., Johansson, O. & Terenius, L. *Neuroscience* **1**, 349–353 (1976).
10. Simantov, R., Kuhar, M. J., Uhl, G. R. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2167–2171 (1977).

11. Ojeda, S. R., Harms, P. G. & McCann, S. M. *Endocrinology* **95**, 1694–1703 (1974).
12. Rivier, C., Vale, W., Ling, N., Brown, M. & Guillemin, R. *Endocrinology* **100**, 238–241 (1977).
13. Lien, E. L., Fenichel, R. L., Garsky, U., Sarantakis, D. & Grant, N. H. *Life Sci.* **19**, 837–840 (1976).
14. Dupont, A., Cusan, L., Garon, M., Labrie, F. & Li, C. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 358–359 (1977).
15. Akil, H., Mayer, D. J. & Liebeskind, J. C. *Science* **191**, 961–962 (1976).
16. Mayer, D. J. *Neurosci. Res. Program Bull.* **13**, 94 (1975).
17. Jacobs, J. J., Tremblay, E. C. & Colombel, M. C. *Psychopharmacology*, **37**, 217–223 (1974).
18. Madden, J., Akil, H., Patrick, R. L. & Barchas, J. D. *Nature* **265**, 358–360 (1977).
19. Shaar, C. J., Frederickson, R. C. A., Dininger, N. B., Clemens, J. A. & Hull, R. H. *Fedn Proc.* **36**, 311 (1977).
20. Lazarus, L. H., Ling, W. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2156–2159 (1976).
21. Bradburg, A. F., Smyth, D. G., Shepp, C. R., Birdsall, N. J. M. & Hulme, E. C. *Nature* **206**, 793–794 (1976).
22. Li, C. H. & Chung, D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1145–1148 (1976).
23. Dupont, A. *et al.* *Fedn Proc.* **36**, 311 (1977).
24. Ferland, L., Fuxe, K., Eneroth, P., Gustafsson, J.-A. & Skett, P. *Eur. J. Pharmacol.* **43**, 89–90 (1977).
25. Halasz, B. & Pupp, L. *Endocrinology* **77**, 553–562 (1965).
26. König, J. F. R. & Klippel, R. A. *The Rat Brain* (Krieger, New York, 1967).

Immunoreactive component resembling cholecystokinin octapeptide in intestine

MANY hormones exist in several molecular forms. The intestinal hormone cholecystokinin has been isolated from hog duodenum in the form of peptides of 33 and 39 amino acid residues (CCK33 and CCK39)^{1–4}. Extracts of hog cerebral cortex have been shown to contain components with CCK-like immunoreactivity⁵, but on gel filtration the main component in brain had the properties of the COOH-terminal octapeptide (CCK8) of CCK33 and CCK39⁵. Several other hormonal and related peptides probably occur in identical forms in brain and gut⁶, so it seemed possible that a CCK8-like factor would also be found in intestine. Synthetic CCK8 is known to be about twice as potent on a molar basis as CCK33 on the main target organs of the hormone (gall bladder and exocrine pancreas)⁷ so that the possible existence of CCK8 in the intestine is of physiological significance. We have studied immunoreactive CCK-like factors in extracts of hog duodenum and jejunum and we report here the presence of relatively high concentrations of a CCK8-like component. This factor was poorly extracted by the method used to purify CCK33 and CCK39 which may explain why it was previously overlooked.

Mutt and Jorpes obtained CCK33 and CCK39 from hog intestine by first immersing the tissues in boiling water to inactivate proteases, and then transferring them to acetic acid for extraction^{1,2}. Basic peptides in the extract, including CCK33 and CCK39, were adsorbed on alginic acid and eluted with HCl. In the present study pieces of hog duodenum, jejunum, antral mucosa and cerebral cortex were extracted either in boiling water or in boiling 0.5 M acetic acid. The tissues (0.5 g ml⁻¹) were briefly boiled (2–3 min), homogenised and centrifuged (2,000g, 10 min). When acid extracts of duodenum or jejunum were fractionated on Sephadex G-50 the main peak of immunoreactivity in the column eluates detected by radioimmunoassay using an antiserum which cross-reacts with CCK8 and CCK33, emerged in the position of CCK33 (Fig. 1, component II^B). This component could be adsorbed on alginic acid and eluted with HCl and is therefore compatible with CCK33 obtained by Mutt and Jorpes. On gel filtration, boiling water extracts of intestine contained little or no activity resembling CCK33 but did contain two other major peaks of immunoreactivity (Fig. 1, components III and IV), which would presumably be discarded in the first step of the Mutt and Jorpes method. The larger of the two peaks (IV) eluted in the region of CCK8 and the other emerged between calibration standards of CCK8 and CCK33 and is therefore probably of intermediate size. Both factors could be ad-

Table 1 Concentration (pmol ml⁻¹) of intestinal peaks III and IV in samples pooled from Sephadex G-50 eluates and estimated by four COOH-terminal specific antisera and pure human G17 or CCK8 standards

	Standard	C	Antiserum		
			2716	1296	L2
Component III	G17	60.00	7.60	2.90	0.70
	CCK8	18.30	47.00	30.00	63.00
Component IV	G17	9.80	0.80	0.40	0.05
	CCK8	3.20	2.60	3.20	4.40

Antiserum C was used with ¹²⁵I-labelled CCK8 when CCK8 was the standard, and with ¹²⁵I-labelled G17 when G17 was the standard. The other antisera were raised against G17 and used with ¹²⁵I-labelled G17 for both G17 and CCK8 standards. Dilution curves of the brain extracts were parallel to the standard curve in each case.

sorbed on DEAE and eluted with ammonium carbonate suggesting that they are acidic peptides. Minor components (I and II^A) were present in both acid and boiling water extracts (Fig. 1). Intestinal components I and IV emerged in the same position as factors previously identified in brain extracts³. Immunoreactivity corresponding to intestinal components II and III was present in negligible amounts in brain, however (Fig. 1). The failure to find significant quantities of a CCK33-like component in brain is not likely to be due to degradation during extraction for standard CCK33 was recovered in good yield (> 80%) when added to acid extracts of brain. Similarly the different patterns of components in acid and boiling water extracts of intestine are not likely to arise from interconversion during extraction, because when the residue of boiling water extracts was re-extracted with 0.5 M acetic acid CCK33-like activity was obtained in good yield (> 80%) compared to direct acid extraction.

Gastrin and CCK share the same COOH-terminal pentapeptide and antisera specific for the COOH-terminus of one hormone frequently cross-react with the other. But, immunochemical studies with antisera of different specificity support the suggestion that intestinal components III and IV are COOH-terminal fragments of CCK. Gastrin antisera (1295 and L6) which do not cross-react with CCK^{8,9}, revealed negligible amounts of immunoreactive material in extracts of duodenum, jejunum or brain (< 0.5 pmol g⁻¹), but high concentrations in antral mucosa (> 2 nmol g⁻¹). Furthermore, in radioimmunoassays using four COOH-terminal specific antisera which cross-react to different degrees with heptadecapeptide gastrin (G17) and CCK8, there were only minor differences between antisera (two- to fourfold) in estimating concentrations of components III and IV when referred to a standard of CCK8 (Table 1). In contrast, when G17 was used as a standard there were up to 200-fold differences in estimates made with different antisera. Similar results have been obtained using a variety of COOH-terminal fragments of gastrin and CCK as standards; together the data indicate that with this panel of antisera the pattern of cross-reactivity of the intestinal components is readily distinguishable from gastrin-like but not CCK-like peptides.

Intestinal components II and III could also be distinguished from G17 and smaller gastrin-like peptides because they were converted by trypsin to a component emerging just before CCK8 (90% elution volume) on Sephadex G-50 whereas G17 and small fragments are resistant to trypsin. A similar tryptic peptide was obtained by trypsinisation of pure CCK33, and probably corresponds to the COOH-terminal dodecapeptide previously identified as a product of mild trypsinisation of CCK33^{2,3}. Intestinal component IV, like CCK8, was resistant to trypsin.

Estimation of the absolute concentration of the different intestinal components must await their purification and

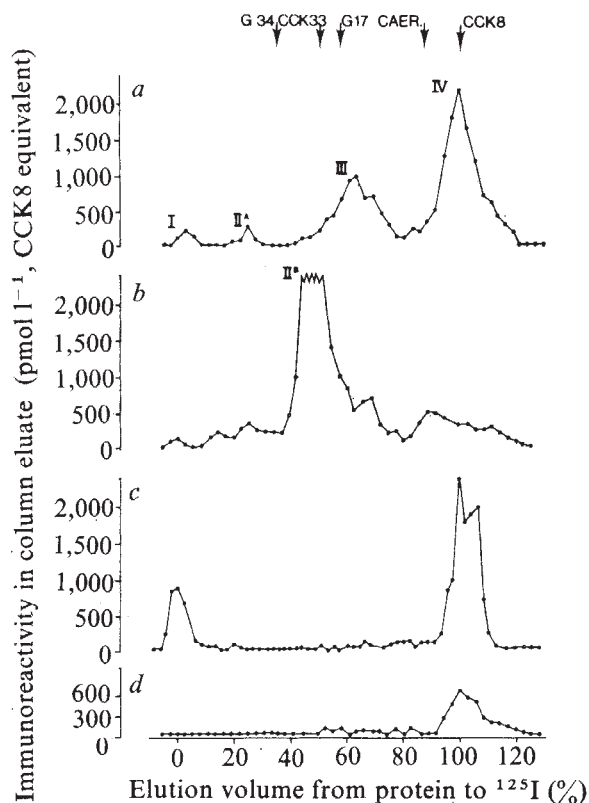


Fig. 1 Elution profiles of boiling water (neutral) and boiling 0.5 M acetic acid extracts of hog jejunum and cerebral cortex on Sephadex G-50 (1 × 100 cm). In each case the equivalent of 0.5 g tissue was applied to the column. Immunoreactivity in the column eluates was estimated by radioimmunoassay using an antiserum ('C') raised against CCK8 and ¹²⁵I-labelled CCK8; with this antiserum the ratio of immunoreactivity CCK8:CCK33, was 1.0:0.60 on a molar basis; immunoreactivity in column eluates is expressed in terms of standard CCK8. Elution volumes are expressed as % from protein peak (0%) estimated by absorption of 280 nm, to ¹²⁵I (100%) added as a marker to all samples. In separate runs the columns were calibrated with pure natural human big gastrin (G34), little gastrin (G17), pure natural porcine CCK33 and CCK39, synthetic caerulein (Caer) and CCK8. Elution volumes of these standards are indicated by arrows; CCK 39 emerges in the same position as CCK 33. a, Intestine neutral extract; b, intestine acid extract; c, brain neutral extract; d, brain acid extract.

characterisation. But, in radioimmunoassays using antisera raised against CCK8, and the appropriate standard (CCK8 or CCK33), it was estimated from immunoreactivity in column eluates that concentrations of the CCK8-like factor in duodenum (40 pmol g^{-1}) were about twice those of CCK33 (19 pmol g^{-1}). In jejunum, on the other hand, concentrations of the CCK8-like factor (15 pmol g^{-1}) were about half those of CCK33 (27 pmol g^{-1}). Using CCK8 standard the concentration of component III was estimated to be 20 pmol g^{-1} in duodenum and 10 pmol g^{-1} in jejunum. Because COOH-terminal fragments of CCK larger than the heptapeptide have strong CCK-like actions⁷, the intestinal CCK8-like factor could make an important, possibly even the major, contribution to the biological activity of cholecystokinin released from the duodenum.

Several interesting questions are posed by the different patterns of distribution of CCK-like peptides in brain and gut. It is now recognised that many polypeptide hormones are synthesised via large precursor molecules which are converted to smaller active molecules by proteolytic cleavage¹⁰. Thus the different forms of CCK could be derived from precursor which was cleaved at different positions in brain and intestine. Even within the small bowel there may be differences in the patterns of biosynthetic processing of CCK, for there were differences in the relative amounts of CCK8- and CCK33-like activity in the duodenum and jejunum. For many hormones, for example, insulin and parathyroid hormone, the site at which the prohormone is cleaved to yield the active hormonal peptide consists of two or more consecutive basic amino acid residues, and is susceptible to digestion by trypsin-like enzymes¹⁰. Only a single basic residue (arginine) links CCK8 to the NH₂-terminal portion of CCK33, however, and this position is relatively resistant to trypsin. If, as seems likely, CCK33 is a prohormone for CCK8 then conversion is probably mediated by a specific enzyme.

We thank Elaine Clarke, Carol Higgins and Gavin Laing for technical assistance, Drs S. Bloom and J. Rehfeld for antisera ('C' and 2716 respectively), Professor R. A. Gregory and Dr H. J. Tracy for pure human G17 and G34, Professor V. Mutt for pure porcine CCK33 and CCK39 and the Squibb Institute of Medical Research for CCK8. The work was supported by a grant from the MRC.

G. J. DOCKRAY

Physiological Laboratory,
University of Liverpool,
Liverpool, UK

Received 2 June; accepted 15 September 1977.

1. Mutt, V. *Arkiv f. Kemi*, **15**, 69–74 (1959).
2. Mutt, V. & Jorpes, E. in *Secretin, Cholecystokinin, Pancreozymin and Gastrin*, 1–179 (Springer Berlin, 1973).
3. Mutt, V. & Jorpes, E. *Eur. J. Biochem.*, **6**, 156–162 (1968).
4. Mutt, V. *Clin. Endocrinol.*, **5**, suppl. 175s–183s (1976).
5. Dockray, G. J. *Nature*, **264**, 568–570 (1976).
6. Pearse, A. G. E. *Nature*, **262**, 92–94 (1976).
7. Grossman, M. I. in *Peptide Hormones* (ed. Parsons, J. A.) 106–116 (Macmillan, London, 1976).
8. Dockray, G. J. & Walsh, J. H. *Gastroenterology*, **68**, 222–230 (1975).
9. Dockray, G. J. & Taylor, I. L. *Gastroenterology*, **71**, 971–977 (1976).
10. Steiner, D. F. in *Peptide Hormones* (ed. Parsons, J. A.) 49–64 (Macmillan, London, 1976).

Specific binding of ³H-mepyramine to histamine H₁ receptors in intestinal smooth muscle

THE specific binding of high-affinity ligands to a wide range of pharmacological receptors has been used to obtain information on receptor numbers, distribution and the characteristics of agonist and antagonist binding. The first study of this type, by Paton and Rang¹, was of the binding of ³H-atropine to muscarinic receptors in guinea pig intestinal smooth muscle. The histamine H₁ receptor has not previously been studied,

and here we describe the specific binding of mepyramine, an H₁-selective antagonist², to histamine H₁ receptors in homogenates of the longitudinal muscle-myenteric plexus preparation from guinea pig small intestine.

³H-Mepyramine was synthesised by catalytic hydrogenation with tritium gas of the 5-bromopyridyl derivative. The method will be described in detail elsewhere. The purity of the product was established by the identity with authentic non-radioactive mepyramine of (1) the *R_F* value on thin-layer chromatography in three solvent systems, (2) the mobilities on high-voltage electrophoresis and (3) the ultraviolet spectrum. The affinity constant of ³H-mepyramine for the histamine H₁ receptor, $1.6 \pm 0.4 \times 10^9 \text{ M}^{-1}$ (three measurements), deduced from the parallel shift of the log dose-response (contraction) curve of longitudinal muscle strips to histamine, was not significantly different from that of non-radioactive mepyramine, $1.2 \pm 0.3 \times 10^9 \text{ M}^{-1}$ (six measurements), the figures being in accord with values in the literature^{2–4}. The concentration of ³H-mepyramine was determined from the absorbance at 310 nm. The specific activity of the purified product was 20 Ci mmol^{-1} .

The binding of ³H-mepyramine, either alone or in the presence of $2 \times 10^{-6} \text{ M}$ non-radioactive mepyramine, to homogenates of longitudinal muscle strips from guinea pig small intestine and the difference curve, which is taken to represent the receptor-specific binding, is shown for a representative experiment in Fig. 1. The difference between all pairs of points was statistically significant ($P < 0.05$). The best-fit⁵ parameters for the experiment in Fig. 1, assuming binding follows simple mass-action kinetics, were $1.7 \pm 1.0 \text{ nM}$ for *K_d*, the dissociation

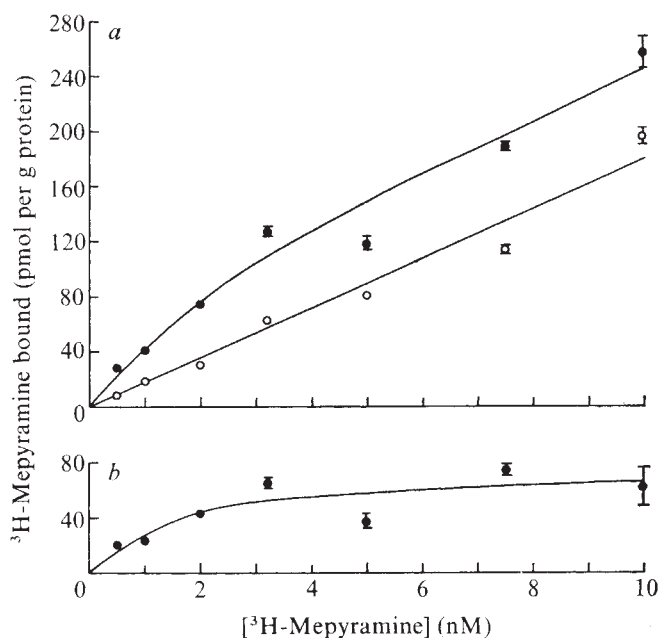


Fig. 1 *a*, Binding of ³H-mepyramine to homogenates of guinea pig intestinal smooth muscle. Longitudinal muscle strips from guinea pig small intestine, prepared essentially as described by Rang¹⁰, were chopped with scissors, partially homogenised in five volumes of 50 mM sodium-potassium phosphate buffer, pH 7.5, in a glass homogeniser with a glass pestle and the ice-cold suspension then treated twice with a Polytron blender (setting 5 for 15 s). Homogenate (0.2 ml, approximately 2 mg protein) in 1.8 ml 50 mM sodium-potassium phosphate buffer, pH 7.5, was incubated with various concentrations of ³H-mepyramine in the presence (○) or absence (●) of $2 \times 10^{-6} \text{ M}$ mepyramine for 60 min at 30 °C. Aliquots (0.25 ml) were centrifuged at 8,700g for 60 s in a Beckman Microfuge B and the pellet washed superficially twice with 0.1 ml ice-cold phosphate buffer. The tritium in the pellet was determined by liquid scintillation counting. Values are means $\pm$ s.e. of quadruplicate determination of two to four incubations. *b*, Receptor-specific binding; obtained by subtracting the binding of ³H-mepyramine in the presence of $2 \times 10^{-6} \text{ M}$ mepyramine from the uninhibited binding. The curve is a best-fit curve calculated by the method of Wilkinson⁵.

constant and 77 ± 15 pmol per g protein for the number of binding sites. The best-fit values obtained by combining data from a number of such experiments were 1.3 ± 0.3 nM and 68 ± 7 pmol per g protein, although the maximum amount of ^3H -mepyramine binding seemed to vary between homogenates (compare Table 2). The affinity constant, $7.7 \pm 1.8 \times 10^8 \text{ M}^{-1}$ ($1/K_d$) is in reasonable agreement with the value, $1.6 \pm 0.4 \times 10^9 \text{ M}^{-1}$, obtained from inhibition of the contractile response.

To establish that the binding of ^3H -mepyramine sensitive to inhibition by 2×10^{-6} M mepyramine does represent receptor-specific binding we have examined the inhibition of ^3H -mepyramine binding by three compounds considered to be selective antagonists at histamine H_1 receptors, chlorpheniramine, promethazine and triprolidine, and by two drugs having their primary effect on other receptors, burimamide (histamine H_2) and lachesine (muscarinic). The maximal inhibition of ^3H -mepyramine binding by each drug did not differ significantly from that produced, measured on the same homogenate, by 2×10^{-6} M mepyramine. The curves for the inhibition by the H_1 antagonists of the receptor-specific binding approximated to hyperbolae and the affinity constants (Table 1) calculated from values for 50% inhibition compared well with those in the literature obtained from inhibition of the contractile response. The two non- H_1 antagonists were only weak inhibitors of ^3H -mepyramine binding. The affinity constant deduced for lachesine, $1.6 \times 10^5 \text{ M}^{-1}$, is in good agreement with the value⁶ from antagonism of the response to histamine, $1.0 \times 10^5 \text{ M}^{-1}$,

Table 1 Inhibition of ^3H -mepyramine binding

Drug	Affinity constant (M^{-1}) from inhibition of		(ref.)
	^3H -mepyramine binding*	Contraction†	
Mepyramine	1.0×10^9	1.2×10^9	
($\pm$) Chlorpheniramine	7.6×10^8	6.6×10^8	(3)
Promethazine	4.4×10^8	8.5×10^8	(3)
Triprolidine	1.0×10^9	c. 10^9	(9)
Burimamide	c. 10^4	3.5×10^3	(8)
Lachesine	1.6×10^5	1.0×10^5	(6)

* Incubations as in Fig. 1 with 10^{-9} M ^3H -mepyramine. The affinity constants, K_a , of the inhibitors were calculated from the relationship $K_a = (K_{mcp} [M] + 1)/[A]$, where $[A]$ is the concentration of drug required for 50% inhibition of receptor-specific ^3H -mepyramine binding, $[M]$ is the concentration of ^3H -mepyramine and K_{mcp} its affinity. K_{mcp} was taken to be $1.2 \times 10^9 \text{ M}^{-1}$, the mean of the various determinations made.

† From inhibition of the contractile response to histamine of intact segments or longitudinal muscle strips from guinea pig ileum. In our own measurements K_a was obtained from the relationship for competitive inhibition: dose ratio $1 = [A] \times K_a$, where the dose ratio is the ratio of the dose of agonist necessary to achieve a given response in the presence of antagonist, concentration $[A]$, to the dose required with no antagonist present.

a figure four orders of magnitude lower than its affinity for the muscarinic receptor⁷, $1.1 \times 10^9 \text{ M}^{-1}$. The affinity of burimamide as an inhibitor of ^3H -mepyramine binding, about 10^4 M^{-1} , is an order of magnitude less than its affinity at the H_2 receptor⁸, $1.3 \times 10^6 \text{ M}^{-1}$.

The number of histamine H_1 receptors in homogenates of guinea pig intestinal smooth muscle, 68 pmol per g protein, is apparently much smaller than the number of muscarinic receptors reported in this tissue¹¹⁻¹³, but the criterion per g protein clearly makes the number obtained very dependent on the degree of purification of the membrane fraction and the assay technique employed. To make a more direct comparison between the numbers of histamine H_1 and muscarinic receptors, in five experiments the maximum receptor-specific binding of ^3H -mepyramine was compared on the same homogenate with that of ^3H -quinuclidinyl benzilate (^3H -QNB), a muscarinic ligand¹⁴. The relative amounts of the two receptors varied within a fairly narrow range (Table 2) giving a mean value of

Table 2 Comparison of the numbers of histamine H_1 and muscarinic receptors in guinea pig intestinal smooth muscle

Expt	No. of receptors (pmol per g protein)		$\text{H}_1/\text{Muscarinic} (\%)$
	Histamine H_1	Muscarinic	
1	82*	496	17
2	41*	407*	10
3	62*	297	21
4	93	493	19
5	45	334	13
Mean	65 ± 10	405 ± 41	16 ± 2

The no. of receptors was estimated either (*) from the maximum specific binding obtained from binding curves in the presence and absence of 2×10^{-6} M mepyramine (for ^3H -mepyramine; see Fig 1) or 2×10^{-6} M methylatropinium bromide (for ^3H -QNB) or (values not starred) from the receptor-specific binding with 7.5×10^{-9} M ^3H -mepyramine or 7.7×10^{-9} M ^3H -QNB. In each experiment measurement of the numbers of the two receptors was made on the same homogenate. ^3H -QNB (^3H -quinuclidinyl benzilate), specific activity 13 Ci mmol⁻¹, was obtained from the Radiochemical Centre, Amersham.

16 ± 2 for the number of H_1 receptors as a percentage of the number of muscarinic receptors. This figure must still, however, be viewed with some caution. The composition of the assay medium, 50 mM sodium potassium phosphate, does not correspond to the *in vivo* environment and while the number of muscarinic receptors determined by antagonist binding seems to be insensitive to the ionic composition¹⁵ it remains to be established whether this also holds for mepyramine.

As far as we are aware this is the first report of binding studies on the histamine H_1 receptor using a selective high-affinity antagonist. There has been a report¹⁶ that by using suitable protecting agents the histamine receptor in cat intestinal muscle could be labelled with the relatively nonspecific antagonist ^3H -dibenamine (58.5 mCi mmol⁻¹). The amount of presumed receptor material labelled seems remarkably large, however, being some 200-fold greater than that bound by ^3H -mepyramine in our experiments and indeed an order of magnitude greater than the highest reported estimate of the number of muscarinic receptors in intestinal smooth muscle. Even allowing for species differences there must be some suspicion that the bulk of the dibenamine was bound to non-receptor sites.

These observations suggest that ^3H -mepyramine should be a valuable tool with which to investigate the numbers of histamine receptors in various tissues in various conditions and to study the binding characteristics of H_1 agonists and antagonists.

S.J.H. is the recipient of an SRC research studentship.

S. J. HILL
J. M. YOUNG

Department of Pharmacology,
University of Cambridge

D. H. MARRIAN

Department of Haematological Medicine,
University of Cambridge, Medical School,
Hills Road, Cambridge, UK

Received 26 July; accepted 30 September 1977.

1. Paton, W. D. M. & Rang, H. P. *Proc. R. Soc. B* **163**, 1-44 (1965).
2. Ash, A. S. F. & Schild, H. O. *Br. J. Pharmac. Chemother.* **27**, 427-439 (1966).
3. Marshall, P. B. *Br. J. Pharmac. Chemother.* **10**, 270-278 (1955).
4. Arunlakshana, O. & Schild, H. O. *Br. J. Pharmac. Chemother.* **14**, 48-58 (1959).
5. Wilkinson, G. N. *Biochem. J.* **80**, 324-332 (1961).
6. Reasbeck, P. G. & Young, J. M. *Br. J. Pharmac.* **48**, 148-155 (1973).
7. Paton, W. D. M. & Rang, H. P. *Adv. Drug Res.* **3**, 57-80 (1966).
8. Black, J. W., Duncan, A. M., Durant, C. J., Ganellin, C. R. & Parsons, E. M. *Nature* **236**, 385-390 (1972).
9. Green, A. F. *Br. J. Pharmac. Chemother.* **8**, 171-176 (1953).
10. Rang, H. P. *Br. J. Pharmac. Chemother.* **22**, 356-365 (1964).
11. Fewtrell, C. M. S. & Rang, H. P. in *Drug Receptors* (ed. Rang, H. P.) 211-224 (Macmillan, London, 1973).
12. Burgen, A. S. V., Hiley, C. R. & Young, J. M. *Br. J. Pharmac.* **50**, 145-151 (1974).

13. Yamamura, H. I. & Snyder, S. H. *Molec. Pharmac.* **10**, 861-867 (1974).
14. Yamamura, H. I. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1725-1729 (1974).
15. Birdsall, N. J. M., Burgen, A. S. V., Hulme, E. C. & Wells, J. W. *Br. J. Pharmac.* **59**, 503P (1977).
16. Takagi, K. & Uchida, M. *Jap. J. Pharmac.* **20**, 272-286 (1970).

Presumptive control mutation for alcohol dehydrogenase in *Drosophila melanogaster*

EXAMPLES of control mutations are rare in eukaryotes. Indeed, the only well-documented examples are in *Aspergillus*^{1,2} and a mutant affecting xanthine dehydrogenase (XDH) activity and mapping adjacent to, but separable from, *rosy*, the structural gene for XDH in *Drosophila melanogaster*^{3,4}. Control of alcohol dehydrogenase (ADH), particularly in *Drosophila* has recently been studied. In examining ADH activity in selection lines and in samples from natural populations, a line was found to be *Adh*^S and to have half the activity of a normal slow-migrating allele. To test the hypothesis that the low activity *Adh*^S line contains a control mutation, we have attempted to separate the activity phenotype from the electrophoretic (structural gene) phenotype by recombination. We have succeeded in doing this, and in the process have obtained a very accurate position for *Adh* based on its recombination with closely-linked marker loci.

ADH is a small homodimer with a molecular weight of about 24,000 (R. Ambler, personal communication). The alcohol dehydrogenase (*Adh*) locus maps at position 50.1 on the second chromosome of *D. melanogaster*, and its cytogenetic location is known within two bands (35B2-3) on the polytene chromosome map of salivary glands⁵.

In addition, *Adh* is highly polymorphic in nature⁶⁻⁸. Activity variation is also known⁹⁻¹¹ and we have identified a striking variant in a population collected near Cambridge, UK. A line derived from that population is homozygous for the slow-migrating allele (*Adh*^S) but has only about half the activity of a standard *Adh*^S allele, as measured by the rate at which it acts on the substrate isopropanol. Taking *Adh*^F as a standard with relative activity of 1.0, we found the ratio of activities of *Adh*^F to *Adh*^S to 'low activity *Adh*^S' to be approximately 1.0 : 0.57 : 0.28 (ref. 12). Radial immunodiffusion assays¹² have confirmed that the decreased activity is due to the fact that the allele produces about half as many enzyme molecules as the normal *Adh*^S or *Adh*^F alleles. Thus, it is a good candidate for a control mutation, though, as is the case for some mutations of XDH, this variant could affect the quantity of ADH by changing the primary sequence of the protein, and only protein chemistry can unambiguously rule out this possibility.

Let us represent the low activity allele as *Adh*^S-L, the normal slow-mobility allele as *Adh*^S-H, and the possible fast-mobility alleles as *Adh*^F-L and *Adh*^F-H, respectively. A multiply-marked second chromosome stock carrying *black* (*b*, 2-48.5), *elbow* (*el*, 2-50.0), *reduced-scraggly* (*rd*^S, 2-51.2), *purple* (*pr*, 54.5), and *cinnabar* (*cn*, 2-57.5) was crossed to the *Adh*^S-L stock, which is marked with the third chromosome mutation *veinlet*. The *Adh* locus maps between *el* and *rd*^S¹³ and the *b el rd*^S *pr cn* stock is homozygous for *Adh*^F-H. Heterozygous females were backcrossed to *b el rd*^S *pr cn* males and the progeny were scored for two recombinant classes: *b el* + + + and + + *rd*^S *pr cn*. All other progeny were pooled and counted as non-recombinants for that region.

A total of 122 *b el* + + + and 128 + + *rd*^S *pr cn* recombinant progeny were recovered among 59,738 scored. This allows us to calculate very accurately the map distance between the two morphological markers flanking the *Adh* locus. This distance is estimated to be 0.42 cM, much smaller than the estimate of 1.2 cM normally quoted¹³. Having such close flanking markers makes the recombinational assay of potential control mutants exceptionally accurate and efficient. Indeed, independent evidence suggests that there may be only one gene between *el* and *Adh*^S.

Each + + *rd*^S *pr cn* recombinant male was then mated to females heterozygous for *Df*(2L)64j (ref. 9), which includes the *el* and *Adh* loci, and for the balancer chromosome *Curly of Oster*. Homozygous + + *rd*^S *pr cn* recombinant chromosomes were produced by standard crosses. A total of 35 of the 70 such chromosomes from males were successfully made homozygous. These were then tested for both ADH electrophoretic mobility and for enzyme activity.

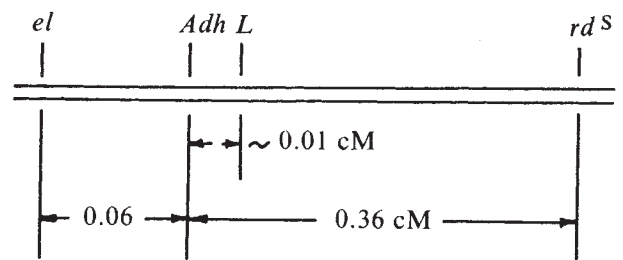


Fig. 1 Order and estimated map distances for *Adh*, the low activity site (L) and flanking markers.

A total of 30 recombinants were *Adh*^S. On the basis of the earlier estimate of the map distance between *el* and *rd*^S, this places the *Adh* locus very near *el*, and about 0.36 cM from *rd*^S (Fig. 1). Of greater importance, however, was the observation that one of the 30 *Adh*^S lines had normal enzyme activity, while the other 29 still had low activity. The complementary *Adh*^F-L recombinant has not yet been recovered. Thus, an activity modifier locus has been identified that is separable from, and to the right of, the *Adh* structural gene as marked by electrophoretic mobility alleles.

Other recombinants are now being studied. But the data so far have enabled us to measure accurately the recombination distance between two loci flanking the *Adh* locus and have provided the first evidence that a *presumptive* control mutation decreasing the number of enzyme molecules is located proximal, but very close, to the structural gene for alcohol dehydrogenase.

This work was supported in part by an MRC grant (G973/691/B) to M.A. and Professor B. Clarke.

JAMES N. THOMPSON, JR

Department of Zoology,
730 Van Vleet Oval,
University of Oklahoma,
Norman, Oklahoma 73019

MICHAEL ASHBURNER

Department of Genetics,
Downing Street,
University of Cambridge,
Cambridge, UK

R. C. WOODRUFF

Department of Biological Sciences,
Bowling Green State University,
Bowling Green, Ohio 43403

Received 2 August; accepted 27 September 1977.

1. Arst, H. N. & MacDonald, D. W. *Nature* **254**, 26-31 (1975).
2. Hynes, M. J. *Nature* **253**, 210-212 (1975).
3. Chovnick, A. *et al. Genetics* **84**, 233-255 (1976).
4. Chovnick, A., Gelbart, W. & McCarron, M. *Cell* **11**, 1-10 (1977).
5. Woodruff, R. C. & Ashburner, M. (in preparation).
6. Johnson, F. M. & Burrows, P. M. *Biochem. Genet.* **14**, 47-58 (1976).
7. Vigue, C. L. & Johnson, F. M. *Biochem. Genet.* **9**, 213-227 (1973).
8. Milkman, R. *Biochem. Genet.* **14**, 383-387 (1976).
9. Grell, E. H., Jacobsen, K. B. & Murphy, J. B. *Ann. N.Y. Acad. Sci.* **151**, 441-455 (1968).
10. Sofer, W. H. & Hatkoff, M. A. *Genetics* **72**, 545-549 (1972).
11. Ward, R. D. & Hebert, P. D. N. *Nature new Biol.* **236**, 243-244 (1972).
12. Thompson, J. N., Jr & Kaiser, T. N. *Heredity* **38**, 191-195 (1977).
13. Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster* (Carnegie Institution of Washington, Publ. No. 627, 1967).

Synthesis of infectious Φ X-174 bacteriophage *in vitro*

STUDY of virus assembly has indicated the involvement of a diverse spectrum of protein-protein and protein-nucleic acid interactions (for review see ref. 1). In *Escherichia coli* cells infected with bacteriophage Φ X-174, synthesis of single-stranded circular viral DNA requires the functions of at least seven phage-coded genes which include genes specifying viral structural proteins F, G and H non-virion proteins A, B, C and D. The only known phage gene that is not involved in single-stranded DNA synthesis is gene E (lysis) (for review see ref. 2). It seems that synthesis of single-stranded circular viral DNA of Φ X-174 is very tightly coupled with the phage morphogenetic pathway. The synthesis of Φ X-174 viral DNA involves a replicative form (RF) DNA with an extended tail of single-stranded viral DNA up to one genome in length ('rolling circle' or σ structure DNA)³⁻⁷. Fujisawa and Hayashi^{8,9} found that σ DNA in infected cells is associated with a number of phage and host proteins as a 50S complex which is a precursor of mature phage particles. They have proposed that synthesis and packaging of Φ X-174 single-stranded DNA occur in the 50S complex. Formation of the 50S complex requires the association of a functional RF DNA template with a 108S capsomeric structure containing the protein products of genes D, F, G and H (ref. 10). The functions of genes B and C are also required for formation of the 50S complex. The gene B protein is necessary to catalyse assembly of gene F and gene G proteins¹¹ to form a precursor of the 108S structure¹⁰⁻¹². In the absence of gene C protein, the 108S structure is produced but the 50S complex is not formed¹⁰. Apparently, RF DNA cannot act as a functional template for formation of the 50S complex and, consequently, for synthesis of single-stranded DNA if the gene C protein is absent. To elucidate the function of Φ X-174-coded proteins during phage morphogenesis and single-stranded DNA synthesis, we have developed an *in vitro* system that synthesises circular, single-stranded viral DNA¹³. The system contained an extract prepared from cells infected with a lysis-defective mutant of Φ X-174. We show here that synthesis of circular viral DNA in the system does depend on the functions of the proteins encoded by Φ X-174 genes B and C. The circular single-stranded viral DNA synthesised *in vitro* is contained in infectious particles whose appearance is very much like that of mature Φ X-174 phage particles when examined by electron microscopy.

To test the requirement for gene B and gene C proteins we performed an *in vitro* complementation experiment. Cell extracts were prepared from *E. coli* H570-22 (*Su*⁻) that had been infected with either a Φ X-174 double mutant (*och11 am3*)¹⁴ of gene C (*ochre*) and gene E (*amber*) or Φ X-174 double mutant (*och12 am3*)¹⁴ of gene C (*ochre*) and gene E (*amber*). The amber mutation in the lysis gene E was present for technical reasons and has no effect on phage DNA replication or morphogenesis. *In vitro* complementation can be performed by mixing infected cell extracts directly in the *in vitro* system but the complementary activity of such a system seems to be limited by a deficiency of phage-specific proteins. In order to increase the activity of the *in vitro* reaction, a fraction (fraction 1) enriched in Φ X-174-specific proteins was prepared from infected cell extracts by precipitation with (NH₄)₂SO₄. *In vitro* reaction mixtures were prepared using the four possible combinations of cell extract and fraction 1, incubated with radioactive dTTP and analysed by sucrose gradient centrifugation. The two reaction mixtures containing homologous combinations of extract and fraction 1 incorporated dTTP only into material sedimenting at 20S (Fig. 1c, d). The 20S material consisted of only RF DNA¹³. No single-stranded DNA was synthesised in these two reaction mixtures which lack either gene B protein or gene C protein. In contrast, the two reaction mixtures containing heterologous

combinations of extract and fraction 1 incorporated dTTP not only in material sedimenting at 20S but also at 50S and 110S (Fig. 1a, b). Sedimentation analysis, in an alkaline sucrose gradient, of DNA isolated from the 110S material showed that the DNA consisted almost entirely of single-stranded circular DNA of unit-genome size (data not shown).

To determine infectivity associated with 110S material, DNA in the 110S material was density-labelled. A reaction mixture containing extracts and fraction 1 was prepared from both *Och11am3* infected cells and *Och12am3* infected cells and was incubated with BrdUTP substituted for dTTP. Figure 2a shows analysis by sucrose-gradient sedimentation of the products.

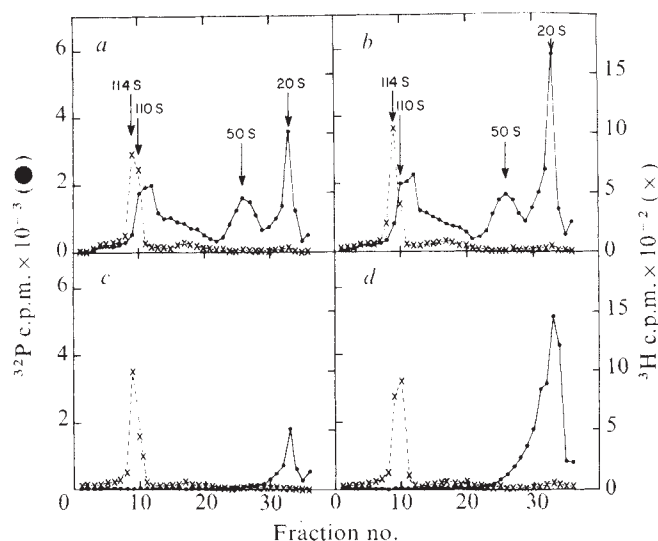


Fig. 1 Sedimentation analysis of *in vitro* product by sucrose gradient centrifugation. *E. coli* H570-22 was infected with Φ X-174 *Och11am3* or with *Och12am3*. Infected cells were collected and washed as described previously¹³. The washed cells were suspended in buffer A (50 mM Tris-HCl, 0.1 mM EDTA, pH 7.3, at 25 °C) containing 10% sucrose and 100 mM KCl and cell extracts were prepared. Saturated (NH₄)₂SO₄ (in 50 mM Tris-HCl, 1 mM EDTA, pH 7.5) was added to the cell extract (0.8 ml) to 40% saturation, and the mixture was left at 0 °C for 30 min. The precipitate was collected by centrifugation at 18,000g for 20 min. The pellet was suspended in 0.1 ml of buffer B (20 mM Tris-HCl, 1 mM MgCl₂, 1 mM dithiothreitol, 10% sucrose, 50 mM KCl, pH 7.5 at 0 °C) and dialysed for 3 h against three changes of 125 ml of buffer B. The dialysed fraction is termed fraction 1. Fraction 1 contained about 20 mg of protein per ml. A standard reaction mixture (0.25 ml) contained 0.14 ml of cell extract, 20 μ l of fraction 1, 50 mM Tris-HCl, pH 7.1 at 30 °C, 0.1 mM each dATP, dCTP, dGTP and α -³²P-dTTP (150–500 c.p.m. pmol⁻¹), 0.1 mM each CTP, UTP, GTP, 0.2 mM ATP, 0.2 mM NAD, 15 mM MgCl₂, 10 mM β -mercaptoethanol, 500 μ g ml⁻¹ bovine serum albumin (BSA) and 80 μ g ml⁻¹ rifampicin. Reaction mixtures were incubated at 30 °C for 18 min. A complete reaction mixture (in 62.5 μ l) (α -³²P-dTTP 220 c.p.m. pmol⁻¹) containing a, 35 μ l of extract from *Och11am3* infected cells and 10 μ l of fraction 1 from *Och12am3* infected cells; b, 35 μ l of extract from *Och12am3* infected cells and 10 μ l of fraction 1 from *Och11am3* infected cells; c, 35 μ l of the extract, 10 μ l of fraction 1 both from *Och11am3* infected cells; or d, 35 μ l of the extract and 10 μ l of fraction 1 both from *Och12am3* infected cells was incubated at 30 °C for 18 min. The reaction was stopped by the addition of 150 μ l of ice-cold buffer C (50 mM Tris-HCl, 100 mM KCl, 15 mM EDTA and 500 μ g ml⁻¹ BSA, pH 7.5 at 0 °C). ³H-Thymidine-labelled mature phage was added as a marker and 200 μ l of the diluted reaction mixture was layered onto a sucrose gradient (15–30% in 10 mM Tris-HCl, pH 7.25 at 25 °C, 100 mM NaCl, 5 mM EDTA and 3.5% CsCl with a 0.15 ml cushion of 55% CsCl and 50% sucrose at the bottom of the gradient). Sucrose gradient was centrifuged in Beckman SW50.1 rotor at 49,000 r.p.m. for 70 min at 0 °C. Fractions were collected onto squares of Whatman 3 MM filter paper which were washed with trichloroacetic acid (TCA) and ethanol, dried and counted. Fractions 37 to 40 were discarded because these fractions did not contain TCA-insoluble material. ●, ³²P c.p.m.; ×, ³H c.p.m.

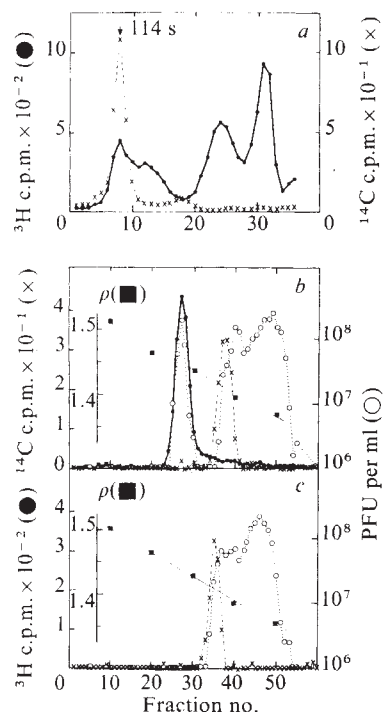


Fig. 2 Analysis of *in vitro* products labelled with BrdUTP. *a*, A reaction mixture (125 μ l) containing 35 μ l of extract and 10 μ l of fraction 1 from *Ochl11am3* infected cells and 35 μ l of extract and 10 μ l of fraction 1 from *Ochl2am3* infected cells was incubated for 18 min at 30 °C. The other components of the reaction mixture were as described in the legend to Fig. 1 except that dTTP and dATP were replaced by 0.1 mM BrdUTP and 0.1 mM 3 H-dATP (440 c.p.m. pmol $^{-1}$) respectively. The reaction was stopped by the addition of 0.5 ml of buffer C. Ultraviolet-irradiated 14 C-marker phage was added to the reaction mixture which was analysed by sucrose gradient centrifugation (5 to 30%, ref. 13) at 49,000 r.p.m. at 0 °C for 50 min. $\bullet$, 3 H; $\times$, 14 C. *b*, Fractions 5 to 8 of the sucrose gradient shown in (*a*) were pooled and 2.2 g of CsCl and 3.1 ml of Na borate buffer (50 mM sodium tetraborate, 5 mM EDTA) were added to 1.05 ml of the pooled material which was dialysed against the borate buffer. Centrifugation was performed at 33,000 r.p.m. for 44 h at 10 °C. Fractions were collected from the bottom of the centrifuge tubes. 10 μ l of each fraction was assayed for radioactivity. The infectivity of each fraction was determined by plaque assay using *E. coli* Su20c (*Su* $^+$) as indicator. $\bullet$, 3 H; $\times$, 14 C; $\circ$, infectivity (PFU per ml); $\blacksquare$, density (g ml $^{-1}$). *c*, A reaction mixture (62.5 μ l) containing 35 μ l of extract and 10 μ l of fraction 1 from *Ochl11am3* infected cells and another reaction mixture (62.5 μ l) containing 35 μ l of extract and 10 μ l of fraction 1 from *Ochl2am3* infected cells were prepared as detailed in (*a*), and incubated separately for 18 min at 30 °C. The reactions were stopped by adding 0.25 ml of ice-cold buffer C to each tube. Then, the two reaction mixtures were combined and centrifuged in a sucrose gradient as described in (*a*). Fractions 5 to 8 of the sucrose gradient were pooled and analysed by CsCl buoyant density centrifugation as described in (*b*). $\times$, 14 C-marker phage; $\circ$, infectivity (PFU per ml); $\blacksquare$, density (ρ , g ml $^{-1}$).

The fractions containing density-labelled 110S material (fractions 5 to 8) from the sucrose gradient shown in Fig. 2*a* were combined and subjected to buoyant density centrifugation in CsCl. The 110S material banded in a single peak at a density of 1.45 g ml $^{-1}$ while mature phage particles not containing BrdUMP banded at a density of 1.41 g ml $^{-1}$ (Fig. 2*b*). When each fraction of the CsCl gradient was titred using *E. coli* Su20c (*Su* $^+$) as the indicator, a peak of infectivity which coincided exactly with the peak of radioactivity at the density of 1.45 g ml $^{-1}$ was found. This peak of infectious material synthesised *in vitro* contained particles that were indistinguishable from mature phage particles when examined in an electron microscope (Fig. 3).

Infectious but non-radioactive material was found at densities equal to and less than the density of marker phage particles (Fig. 2*b*). This material was also present in *in vitro* reaction

mixtures in the absence of complementation (Fig. 2*c*). Since no material labelled *in vitro* was seen at these densities, the low density infectious material apparently existed in the infected cell extracts and might be due to unclipped phages that were recovered from infected cells in the extracts and/or fraction 1.

Several observations indicate that synthesis and packaging of Φ X-174 single-stranded DNA by our *in vitro* system proceed by mechanisms very similar to those that occur *in vivo*. Our previous results suggest that single-stranded DNA synthesis *in vitro* involves the 50S complex containing σ structure DNA 13 and that the *in vitro* 50S complex is a precursor of a particle containing circular single-stranded DNA 13 . We have shown here *de novo* synthesis of single-stranded viral DNA via the 50S complex by complementation experiments (Fig. 1). The synthesised single-stranded DNA is packaged in particles that are infectious and are similar in appearance to mature phage

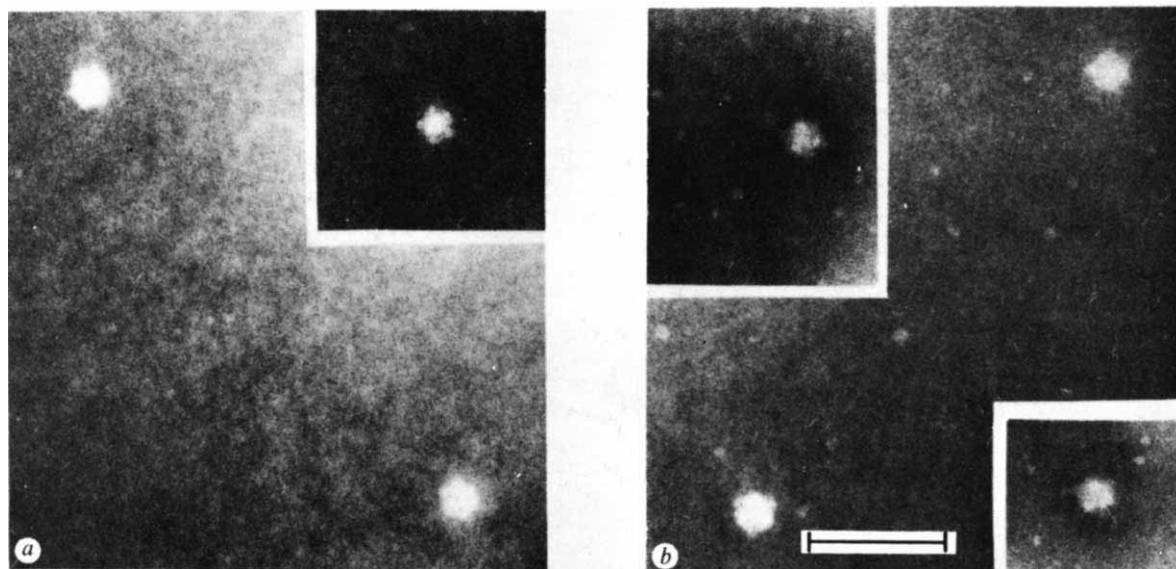


Fig. 3 Electron micrographs of 110S particles. Preparations of mature phages or density-labelled 110S particles (Fig. 2*b*) containing 1–3 $\times 10^{11}$ particles per ml were negatively stained with 2% neutral potassium phosphotungstate. *a*, Mature phages; *b*, 110S particles. Scale bar, 100 nm.

particles in the electron microscope. Further studies are required to determine the differences between 110S *in vitro* particles and 114S *in vivo* mature phages. The *in vitro* complementation system for synthesis of infectious Φ X-174 particles described here may provide a new way of studying the interaction of viral components during phage morphogenesis.

We thank Dr F. Fujimura and Mrs M. Hayashi for discussions, Dr F. D. Funk for providing mutant phages and Mr P. Kammerer for technical assistance. This work was supported by USNSF and NIH grants.

RYOZABURO MUKAI
MASAKI HAYASHI*

Department of Biology,
B-022, University of California,
San Diego, La Jolla, California 92093

Received 16 May; accepted 21 September 1977.

* To whom correspondence should be addressed.

1. Casjins, S. & King, J. A. *Rev. Biochem.* **44**, 555-611 (1975).
2. Denhardt, D. T. *CRC Crit. Rev. Microbiol.* **4**, 161-223 (1975).
3. Dressler, D. *Proc. natn. Acad. Sci. U.S.A.* **67**, 1934-1942 (1970).
4. Dressler, D. H. & Denhardt, D. T. *Nature* **219**, 346-351 (1968).
5. Gilbert, W. & Dressler, D. *Cold Spring Harb. Symp. quant. Biol.* **33**, 473-484 (1968).
6. Knippers, R., Razin, A., Davis, R. & Sinsheimer, R. L. *J. molec. Biol.* **45**, 237-263 (1969).
7. Sinsheimer, R. L., Knippers, R. & Komano, T. *Cold Spring Harb. Symp. quant. Biol.* **33**, 443-448 (1968).
8. Fujisawa, H. & Hayashi, M. *J. Virol.* **19**, 409-415 (1976).
9. Fujisawa, H. & Hayashi, M. *J. Virol.* **19**, 416-424 (1976).
10. Fujisawa, H. & Hayashi, M. *J. Virol.* **21**, 506-515 (1977).
11. Siden, E. J. & Hayashi, M. *J. molec. Biol.* **89**, 1-16 (1974).
12. Tonegawa, S. & Hayashi, M. *J. molec. Biol.* **48**, 219-242 (1970).
13. Mukai, R. & Hayashi, M. *J. Virol.* **22**, 619-625 (1977).
14. Funk, F. D. & Sinsheimer, R. L. *J. Virol.* **6**, 12-19 (1970).

Extensive reverse transcription of RSV genome by nucleic acid-binding protein

THE discovery of reverse transcriptase^{1,2} offered the possibility of using *in vitro* synthesis to study the replication of RNA tumour virus genomes in molecular detail. Attempts to study DNA provirus formation *in vitro*, however, have encountered serious obstacles: poor efficiency of transcription, small DNA products, and unequal representation of the viral RNA sequences in the DNA product have been common findings³. These difficulties in copying RNA into DNA have led to suggestions that cellular factors, such as ligases or unwinding proteins might be necessary to assist reverse transcriptase in copying the viral RNA genome into a complete DNA transcript^{4,5}.

Binding (unwinding) proteins involved in DNA synthesis are well known in bacteria^{6,7} and have also been described in mammalian systems⁸. We reported previously the isolation of a DNA binding protein from transformed chicken cells and its stimulating effect on reverse transcriptase activities⁵. Here we present evidence that, in the presence of the binding protein, purified reverse transcriptase from Rous sarcoma virus (RSV) synthesised an extensive, possibly complete complementary transcript of DNA from the viral RNA genome in a totally reconstituted reaction system.

RSV-transformed chicken fibroblasts were sonicated and its nucleic acids were removed by polyethylene glycol precipitation in high salt as reported previously⁵. The supernatant was then loaded on a single-stranded DNA-cellulose column and the binding activity eluted with 0.2 M NaCl was further passed through DEAE-cellulose and then to a CM BioGel (BioRad) column. A single peak of binding activity as measured by the membrane filtration technique⁹ was eluted from the CM BioGel column at 0.3 M NaCl in a buffer containing 0.05 M Tris-HCl, pH 8.1, 10% glycerol, 1 mM 2-mercaptoethanol, and 1 mM EDTA. This material showed a major Coomassie blue band of over 85% homogeneity corresponding to a molecular weight of about 30,000. One microgram of the purified protein was capable of increased binding to 16 μ g of ³H-poly (dA,dT) and thus at least a 500-fold enrichment in the nucleic acid binding activity was obtained from the crude sonicated extract. It had no detectable reverse transcriptase, DNA polymerase or nuclease activities when analysed¹⁰⁻¹². Detailed isolation procedures will be published elsewhere. The binding protein was able to bind to various synthetic polynucleotides as well as to 70S ³H-RSV RNA. Figure 1a and b shows the binding activity profiles when 70S ³H-RSV RNA and ³H-poly(A,U) were used as the substrates. The amount of binding as indicated by radioactivity retained on the membrane increased linearly until the protein saturated the binding capacity. It was calculated that at the saturation point one binding protein molecule covered 10 nucleotide bases in these substrates. Since the protein showed different rates in binding to various nucleic acids (unpublished observation), we studied the affinity of the protein to a few kinds of nucleic acid in environments of increasing ionic strength. Figure 2 shows the results of the binding study at different NaCl concentrations. There were significant decreases in binding of the protein to ³H-poly A, ³H-poly(A,U) and ³H-poly

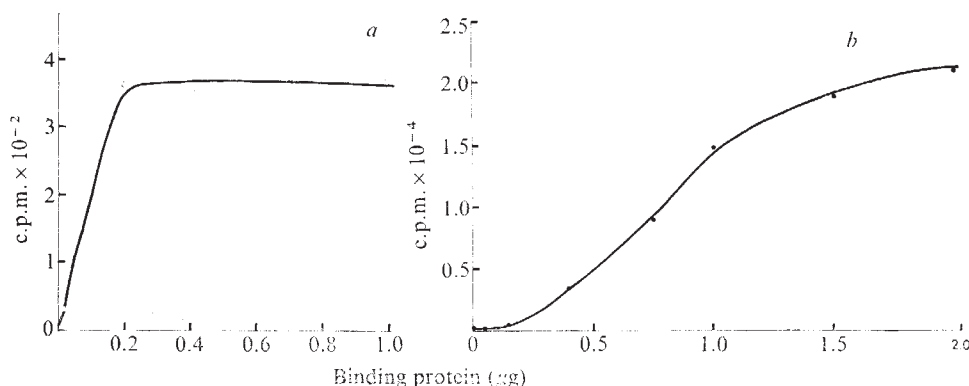


Fig. 1 Nucleic acid-binding activities of the binding protein as a function of its concentration. The assay measured the amount of radioactive nucleic acid bound to nitrocellulose filters⁹ in the presence of the binding protein. The binding buffer contained 0.02 M Tris-HCl (pH 8.1), 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.02 M NaCl and 5% glycerol. After incubation at 37 °C for 10 min in a total volume of 0.1 ml, the reaction mixture was chilled, diluted and filtered through a Millipore nitrocellulose filter (type HA 0.45 μ m). The filter was washed with the binding buffer, dried and the radioactivity determined in a toluene-based scintillation fluid. The binding substrates: a, 70S ³H-RSV RNA (0.2 μ g, 20,000 c.p.m. μ g⁻¹); b, poly(³H-A,U) (0.1 μ Ci; 34 mCi per mmol phosphate).

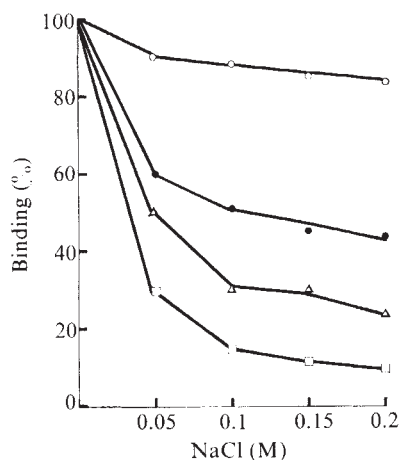


Fig. 2 Effects of ionic strength on the binding of the protein to various nucleic acids. The buffer used during binding and washing was the same as described in Fig. 1, except that the concentration of NaCl varied as indicated. The binding in the absence of NaCl was taken as 100%. ○, 70S ³H-RSV RNA (0.2 μg, 20,000 c.p.m. μg⁻¹); ●, ³H-poly A (0.1 μCi; 25 mCi; per mmol per mmol phosphate); △, poly (³H-A,U) (0.1 μCi; 34 mCi per mmol phosphate); □, ³H-poly(dA, dT) (0.1 μCi; 22 mCi per mmol phosphate). Radioactive synthetic polynucleotides were obtained from Miles.

(dA,dT) when NaCl concentration was increased to 0.05 M or higher. The binding to 70S ³H-RSV RNA by the protein decreased only 10% at 0.2 M NaCl, however. It is possible that some specific nucleic acid sequences in the RSV RNA have a high affinity to the binding protein and thus make the complex more resistant to dissociation in higher ionic strength. An additional piece of evidence to substantiate that the binding protein interacts with nucleic acid was provided by the following study. It is well known that the single-stranded viral RNA contains considerable secondary structures with certain regions of the genome intramolecularly hydrogen-bonded¹³. Therefore we studied the unwinding of this highly structured RNA by the binding protein by measuring ultraviolet hyperchromicity

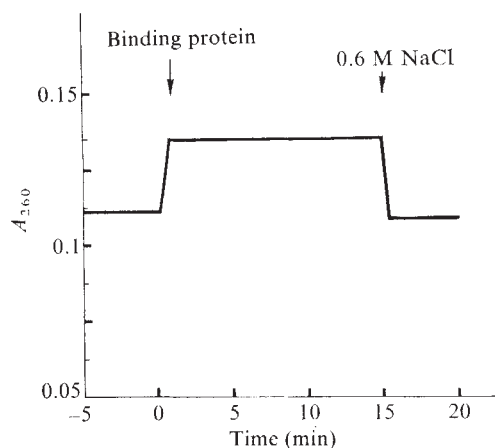


Fig. 3 Unwinding of the 70S RSV RNA by the binding protein. The reaction mixture contained 70S RSV RNA (1.6 μg) and the binding protein (15 μg) in 0.5 ml of 0.05M Tris-HCl (pH 8.1), 0.02 M NaCl, 1 mM 2-mercaptoethanol, 1 mM EDTA and 30% glycerol. Binding protein and the RNA were mixed at time zero, and A_{260} was determined with a recording Gilford Spectrophotometer at 24 °C. Solutions of the binding protein or the RNA alone in the same buffer were scanned at the same time as controls—no changes in absorbance were observed.

which is characteristic of the helix-coil transition¹⁴. The result of such an experiment is shown in Fig. 3. The absorption at 260 nm of 70S RSV RNA increased 15–20% in less than 1 min after the addition of the binding protein at 24 °C and this suggested a fast unwinding of RSV RNA by the binding protein. The hyperchromic effect was reversible on addition of NaCl to 0.6 M (Fig. 3) suggesting that the RNA was deproteinated and returned to the folded structure.

Because of the closeness of the natural primer to the 5' terminus of RSV RNA (refs 15, 16) and of the great secondary structure of the RNA, *in vitro* DNA synthesis can only proceed to a limited distance along the template and result in small DNA products⁸. We therefore attempted to remedy these two difficulties during reverse transcription by, first, using (dT)₁₂₋₁₈ as the primer to start DNA synthesis from the 3' terminus¹⁷ and, second, by adding the binding protein to melt the hydrogen-bonded regions. DNA was synthesised in a totally reconstituted system containing purified 70S RSV RNA (ref. 18), reverse transcriptase¹⁰, dATP, dGTP, dCTP, Mg²⁺, ³²P-TTP, (dT)₁₂₋₁₈ and actinomycin D as described in Fig. 4. After incubation the reaction was stopped by EDTA and passed through a Sephadex G-50 column, DNA in the void volume was ethanol-precipitated, alkaline-hydrolysed and analysed by Agarose gel electrophoresis¹⁹. The results of the experiment are shown by autoradiography in Fig. 4. Lane 1 shows that when no binding protein was added most DNA products moved to the far end of the gel corresponding to a mobility of about 4S.

Table 1 Hybridisation of DNA products with 70S RSV RNA

Sample	70S RSV RNA	Acid-precipitable radioactivity (c.p.m.)	Resistant fraction (%)
Control (minus S ₁ nuclease)	—	1,120	100
Self-annealing (plus S ₁ nuclease)	—	71	6
Hybridisation (plus S ₁ nuclease)	+	1,066	95

Radioactive DNA was prepared as described in Fig. 4, lane 2. DNA samples were hybridised in 2×SSC at 68 °C for 2 h in a total volume of 20 μl. When 70S RSV RNA (2 μg) was used, the ratio of RNA to DNA was estimated to be over 1,000. Resistance to S₁ nuclease was determined by incubating an aliquot of the incubation mixture for 2 h at 37 °C in 1 mM ZnSO₄, 0.1 M sodium acetate (pH 5.0), and 2,000 units of S₁ nuclease, when used. The nuclease-resistant DNA was recovered as trichloroacetic acid-precipitable counts with 10 μg of salmon sperm DNA as carrier.

In lanes 2, 3 and 4 with various preparations of binding protein added, almost all of the DNA product moved much slower and its mobility corresponded to that of 35S RSV RNA. The appearance of the large products in lane 2, 3 and 4 was eliminated if the binding protein solution was first heated to 65 °C for 10 min (not shown here). To confirm that the radioactive product which corresponded to 35S in electrophoretic mobility was indeed complementary to 70S RSV RNA, we annealed the DNA product to the RNA template and digested the unreacted DNA with nuclease S₁ (ref. 20). Table 1 shows that 95% of radioactivity in the DNA product was resistant to the digestion after it had hybridised to 70S RSV RNA whereas DNA by itself was sensitive to the nuclease. This strongly suggested that the product was a single-stranded DNA with homology to the viral RNA genome.

The results presented here show that reverse transcriptase from RSV is able to synthesise, in a reconstructed system, a complete or nearly complete DNA copy from the purified viral RNA in the presence of the binding protein isolated

from chicken cells. There are many advantages in being able to make a complete cDNA in a totally reconstructed system. The long, representative DNA transcripts of tumour viral RNA provide a source of excellent probes for molecular hybridisation. By controlling components of the *in vitro* reaction one at a time, we can obtain a molecular detail of the mechanism of complete reverse transcription.

Although many investigators attempted to synthesise the complete cDNA from tumour virus RNA in a purified enzyme system, the DNA product was always considerably smaller than the template³. In some cases, there was faithful representation of the sequence information in the RNA,

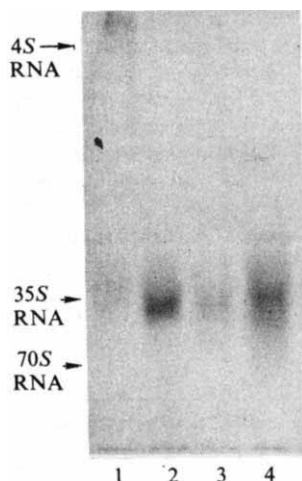


Fig. 4 Autoradiography of DNA products synthesised *in vitro* by reverse transcriptase. Incubation mixtures (0.1 ml) contained the following: 0.05 M Tris-HCl, pH 8.3, 0.06 M KCl, 5 mM MgCl₂, 0.01 M dithiothreitol, 0.1 mM EDTA, 0.02% NP-40, 1.6 µg 70S RSV RNA, 0.1 µg (dT)₁₀₋₁₈, 100 µg ml⁻¹ actinomycin D, 0.1 mM each of dATP, dCTP, dGTP, 10 µCi [α -³²P]TTP (1 Ci mmol⁻¹), 20 units of reverse transcriptase purified by the method of Grandgenett *et al.*¹⁰. The reaction mixture was incubated at 37°C for 10 min, and then binding protein or buffer alone was added. Lane 1, buffer control. Lane 2, 4 µg of binding protein from a preparation of transformed cells. Lane 3, 4 µg from another transformed cell preparation. Lane 4, 15 µg of binding protein from non-transformed cells. Incubation was carried out further at 24°C for 4 h. At the end of the incubation, the reaction mixture was adjusted to contain 0.05 M EDTA, 0.2% sodium sarkosyl and 50 µg yeast RNA. The DNA product was isolated by Sephadex filtration and alkali hydrolysis according to Rothenberg and Baltimore¹⁹. The samples thus obtained were divided into two parts, one for electrophoresis followed by autoradiography and the other, for hybridisation studies. Electrophoresis of the DNA products (approximately 1,000 c.p.m. per slot) was done in 0.8% Agarose gel (10 cm long) for 3.5 h at 100 V (ref. 25). Non-radioactive 70S RSV RNA, 35S RSV RNA obtained by heating and chilling of 70S (ref. 26) and tRNA were co-electrophoresed. Ethidium bromide was used to stain these markers after the electrophoresis and arrows indicate their positions. The gel was dried and autoradiographed as described by Maizel *et al.*²⁷.

but only a small portion of the template was represented in a greatly disproportionated amount^{19,21,22}. The incomplete cDNA transcripts of the RSV RNA occurred as discrete size classes rather than a continuous array of decreasing sizes²³. This suggested that when reverse transcriptase started DNA synthesis at the primer and transcribed along the single-stranded RNA, it could encounter hairpin loops or other forms of secondary structure present in the viral RNA at one or more points, thereby resulting in a partial blockage of transcription or in dissociation of the enzyme at these discrete sites. The nucleic acid-binding protein, by virtue of its ability to bind and unwind the RNA molecule

(Figs 1 and 3), can hold the RNA in a rigid, extended conformation which affords a favourable configuration for reverse transcriptase to continue the cDNA synthesis instead of coming to a stop at the hairpin loops. We believe that the affinity of the binding protein to RSV RNA is such that when reverse transcriptase moves along the template, it replaces the binding protein to continue cDNA synthesis. This was supported by evidence shown in Fig. 5 that reverse

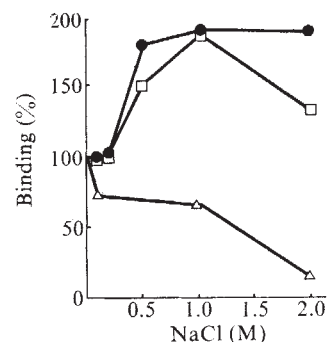


Fig. 5 Effect of ionic strength on the binding of reverse transcriptase to various nucleic acids. The buffer used during binding and washing was the same as described in Fig. 1, except that the concentration of NaCl varied as indicated. Binding of reverse transcriptase (5 units) to the substrate in the absence of NaCl was taken as 100%. ●, 70S ³H-RSV RNA (0.2 µg; 20,000 c.p.m. µg⁻¹); ×, ³H-poly U (0.1 µCi; 30 mCi per mmol phosphate); △, poly(³H-A, U) (0.1 µCi; 34 mCi per mmol phosphate).

transcriptase bound to RSV RNA much tighter than the binding protein. In this experiment, the binding of reverse transcriptase to 70S RSV RNA increased twice at ionic strengths as high as 1.0 M NaCl, in contrast to a gradual dissociation of the binding protein from RSV RNA at higher ionic strength (Fig. 2).

Although the binding protein reported here was purified from RSV-transformed chick cells, the protein did not seem to be coded by the virus. In a preliminary study we also detected a binding protein of similar properties in normal chicken cells (Fig. 4, lane 4). A viral protein binding tightly to the viral RNA has been recently found in the RSV viral core²⁴. This viral core protein formed such a tight complex with the viral RNA that it inhibited rather than stimulated reverse transcription (J. Leis, personal communication).

We are now characterising the DNA product in more detail. It will be of great interest to synthesise a DNA duplex (provirus) of the whole viral genome and study its infectivity as well as nucleotide sequences of the fragments obtained from restriction enzymes.

We thank W. Schulze for technical assistance.

S. G. LEE
P. P. HUNG

Molecular Virology Laboratory,
Abbott Laboratories,
North Chicago, Illinois 60064

Received 1 August; accepted 2 September 1977.

- Baltimore, D. *Nature* **226**, 1209-1211 (1970).
- Temin, H. M. & Mizutani, S. *Nature* **226**, 1211-1213 (1970).
- Green, M. & Gerard, S. F. *Prog. Nucleic Acids Res. molec. Biol.* **14**, 187-334 (1974).
- Temin, H. M. & Baltimore, D. *Adv. Virus Res.* **17**, 129-186 (1972).
- Hung, P. P. & Lee, S. G. *Nature* **259**, 499-502 (1976).
- Sigal, N., Dalius, H., Kornberg, A., Gefter, M. L. & Alberts, B. M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3534-3541 (1972).
- Alberts, B. M. & Frey, L. *Nature* **227**, 1313-1318 (1970).
- Herrick, G. & Alberts, B. W. *J. biol. Chem.* **251**, 2124-2132 (1976).
- Tsai, R. L. & Green, H. J. *molec. Biol.* **73**, 307-316 (1973).
- Grandgenett, D. P., Gerard, G. F. & Green, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 230-234 (1973).
- Apshian, H. V. & Kornberg, A. *J. biol. Chem.* **237**, 516-525 (1962).
- Holloman, W. K. & Holliday, R. *J. biol. Chem.* **248**, 8107-8113 (1973).
- Delius, H., Duesberg, P. H. & Mangel, W. F. *Cold Spring Harb. Symp. quant. Biol.* **39**, 835-843 (1975).

14. Marmur, J. & Doty, P. *Nature* **183**, 1427-1429 (1959).
15. Staskus, K. A., Colletti, M. S. & Faras, A. J. *Virology* **71**, 162-168 (1976).
16. Taylor, J. M. & Illmensee, R. J. *Virol.* **16**, 553-558 (1975).
17. Taylor, J. M. *et al.* *Biochemistry* **12**, 460-467 (1973).
18. Robinson, W. S., Robinson, H. L. & Duesberg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **58**, 825-834 (1967).
19. Rothenberg, E. & Baltimore, D. J. *Virol.* **21**, 168-178 (1977).
20. Ando, T. *Biochim. biophys. Acta* **114**, 158-168 (1966).
21. Junghans, P., Duesberg, P. H. & Knight, C. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4895-4899 (1975).
22. Collett, M. S. & Faras, A. J. J. *Virol.* **16**, 1220-1228 (1975).
23. Haseltine, W. A., Kleid, D. G., Panet, A., Rothenberg, E. & Baltimore, D. *J. molec. Biol.* **106**, 109-131 (1976).
24. Sen, A., Sherr, C. J. & Todaro, G. J. *Cell* **7**, 21-32 (1976).
25. Adesnik, M. *Meth. Virol.* **5**, 125-177 (1971).
26. Canaan, E. & Duesberg, P. J. *Virol.* **10**, 23-31 (1972).
27. Maizel, J. V. Jr. *Meth. Virol.* **5**, 179-245 (1971).

2-Aminoadenine is an adenine substituting for a base in S-2L cyanophage DNA

THE properties of DNA from viruses acting on blue-green algae (cyanophages) have been studied in only a few instances^{1,2}. It is not known whether these DNA molecules contain any unusual bases³⁻⁷ or some modified (methylated) bases known to be present in DNA of many prokaryotes and to be specific for host modification and restriction phenomena. Here we report the isolation and identification of the base 2-aminoadenine and its deoxyribonucleoside from S-2L cyanophage DNA. Some unusual properties of this DNA in which adenine is completely substituted with 2-aminoadenine are described.

Cyanophage S-2L was isolated from water samples taken in the outskirts of Leningrad. It can lyse blue-green algae *Synechococcus* sp. 698 and *S. elongatus* strains 58 and 6907 (from the collection of the Biology Institute, Leningrad State University). The phage was grown on *S. sp.* 698 and concentrated by chromatography on DEAE-cellulose (Whatman DE-32) and purified by centrifugation in CsCl gradient^{2,8}. The phage consists of an icosahedral head, 56 nm in diameter, with a flexible non-contractile tail 120 nm long and 10 nm thick. S-2L phage is morphologically similar to SM-2 cyanophage⁹.

The nucleic acid of phage S-2L is linear double-stranded DNA. It reacts typically for deoxyribose with a diphenylamine-containing reagent according to Dische; it is not hydrolysed in

alkali (0.7 M NaOH, 18 h at 37 °C) or with pancreatic RNase. The mean contour length of S-2L DNA molecules prepared according to Kleinschmidt¹⁰ was $14.4 \pm 0.5 \mu\text{m}$, corresponding to a molecular weight¹¹ of $28 \pm 0.5 \times 10^6$. The average value for $S_{20,w}^{0.0}$ for S-2L DNA was 32.5 ± 0.3 S—a molecular weight value¹² of $26 \pm 1 \times 10^6$.

The buoyant density of S-2L DNA, relative to *Escherichia coli* B DNA ($\rho = 1.7109 \text{ g cm}^{-3}$) was found to be 1.731 g cm^{-3} . When S-2L DNA was centrifuged with *Micrococcus lysodeikticus* DNA ($\rho = 1.731 \text{ g cm}^{-3}$), a single symmetric ultraviolet absorbing peak was observed.

The hyperchromic value of S-2L DNA at 260 nm is 27-29%, at 280 nm it reaches 50% (Fig. 1); T_m in 0.1 SSC is 85.6 °C; in 0.01 M sodium phosphate buffer pH 7.0, 0.001 M Na₂ EDTA T_m is 84.3 °C. The DNA melts within a small temperature range ($\Delta T_m = 3$ °C) which is typical of homogenous viral DNA. A characteristic feature of native S-2L DNA is the unusually low spectral ratio ($A_{\text{max}}/A_{\text{min}} = 1.6$, $A_{260}/A_{230} = 1.5$, $A_{260}/A_{280} = 1.75$).

In contrast to the circular dichroism (CD) spectrum of usual natural DNA that possesses a single positive band at 276 nm, S-2L DNA has two almost equal positive bands, at 290 nm and 265 nm, (Fig. 2). This very characteristic CD spectrum and the other unusual ultraviolet spectral properties strongly suggest the presence of some unusual base in S-2L phage DNA.

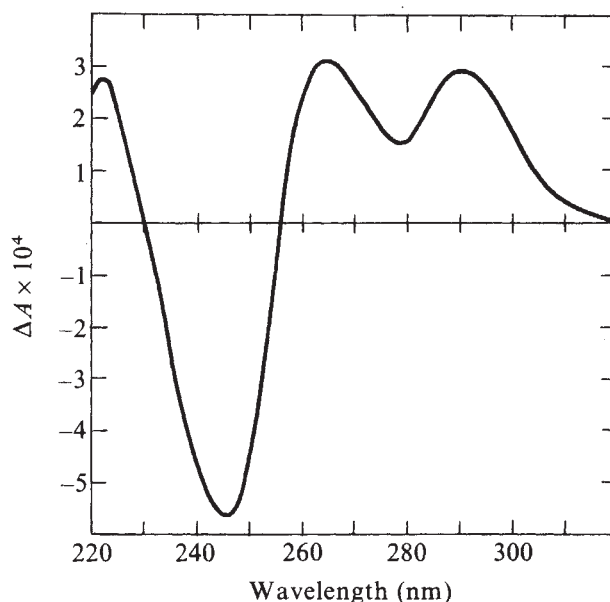


Fig. 2 CD spectrum of S-2L phage DNA in 0.01 M phosphate buffer, pH 7.0-0.001 M Na₂-EDTA (1.3 absorbance units at 260 nm per ml).

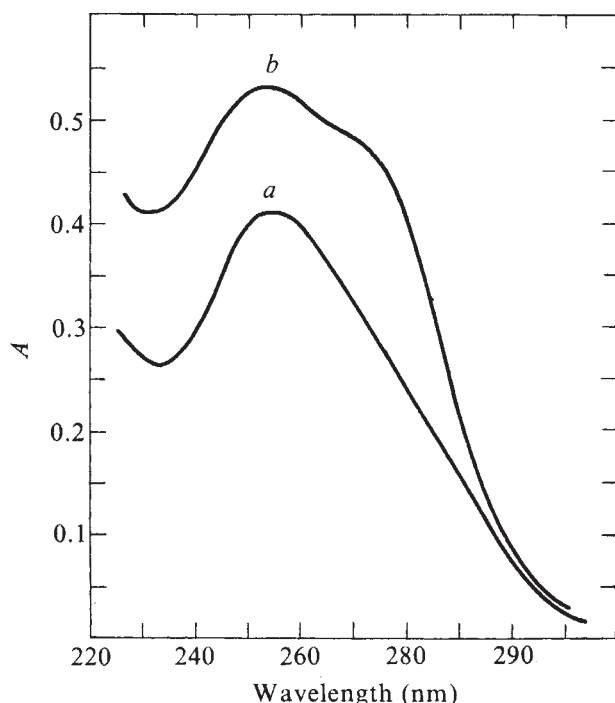


Fig. 1 S-2L phage DNA absorption spectrum in 0.1 SSC; a, at 20 °C and b, at 90 °C.

For determination of the base composition, the DNA was hydrolysed into bases by 57% HClO₄ (100 °C, 1 h) or 99% HCOOH (176 °C, 1 h). Bases were separated by thin layer chromatography (TLC) (Fixion 50 × 8 cation exchanger (Chinoin) in 2.8 M HCl)¹³; as well as thymine, cytosine and guanine, a fourth base with the lowest R_f value, was present. This base was purified by TLC on cellulose (Filtrak) in different solvent systems. The isolated base shows intensive blue fluorescence in ultraviolet light and a characteristic ultraviolet adsorption spectrum (two maxima); the base has the same spectral (Fig. 3, Table 1) and chromatographic properties as authentic 2-aminoadenine (Chemapol). On bromination the base degrades—like 2-aminoadenine and unlike adenine (Fig. 3).

Complete enzymatic hydrolysis of S-2L DNA to nucleosides, yielded 2-aminoadenine deoxyribonucleoside after two-dimensional TLC of deoxyribonucleosides on cellulose¹⁴. The ultraviolet and fluorescence spectra of the 2-aminoadenine deoxyribonucleoside isolated were identical to those reported for

Table 1 Spectral properties of the novel base and deoxyribonucleoside from cyanophage S-2L DNA compared with synthetic 2-aminoadenine

Spectral parameters	Phosphate buffer 0.05 M pH 6.7		Spectral properties in			
	Deoxyribonucleoside*	Base†	0.1 M HCl 2-Aminoadenine treated as base from phage		0.1 M KOH	
$\lambda_{\max}(1)$, (nm)	254	241.5	241.5	241	Inflection (243 nm)	Inflection (243 nm)
$\lambda_{\max}(2)$, (nm)	279	282	282	281.5	283	283
$\lambda_{\min}(1)$, (nm)	233	231.5	231.5	231.5	—	—
$\lambda_{\min}(2)$, (nm)	263	261.5	261.5	261.5	258	257
$A_{\max}(1)/A_{\min}(2)$	0.12	1.75	1.75	1.80	—	—
$A_{\max}(2)/A_{\min}(2)$	0.13	1.86	1.86	1.90	2.46	2.42
A_{230}/A_{260}	0.72	1.45	1.60	1.55	—	—
A_{240}/A_{260}	0.72	1.73	1.80	1.80	1.32	1.26
A_{250}/A_{260}	0.12	1.50	1.50	1.30	1.25	1.10
A_{270}/A_{260}	0.11	1.30	1.30	1.45	1.39	1.55
A_{280}/A_{260}	1.29	1.84	1.85	1.89	2.40	2.38
A_{290}/A_{260}	0.88	1.51	1.50	1.53	2.20	2.10
Excitation maximum in 0.05 M K phosphate buffer (λ nm)	300	300	300	300	—	—
Fluorescence maximum (λ , nm)	360	360	360	—	—	—

*Isolated from hydrolysed phage DNA by two-dimensional TLC on cellulose (Filtrak)¹⁴.

†Isolated by TLC on Fixion 50 × 8 in 2.8 M HCl¹³, repeated TLC on cellulose (Filtrak) in isopropanol–11 M HCl–H₂O (170:45:35 v/v) and then repeated TLC in the solvent systems *n*-butanol–H₂O–25% NH₄OH (60:10:0.1 v/v); isopropanol–H₂O (70:30), NH₃ in gas phase; *n*-butanol–5 M acetic acid (2:1 v/v).

2-aminoadenosine¹⁵. Thus, S-2L phage DNA contains 2-aminoadenine, guanine, cytosine and thymine; no other bases were found. DNA of host alga *Synechococcus* sp. 698 contains usual bases, including adenine, rather than 2-aminoadenine, and its GC content is 70.0 mol%.

On complete enzymatic hydrolysis the S-2L DNA was found to contain (mol %): 2-aminoadenine deoxyribonucleoside, 15.9 ± 0.5; deoxythymidine, 15.4 ± 0.1; deoxycytidine, 34.4 ± 0.4 and deoxyguanosine, 34.3 ± 0.3. This DNA thus contains equimolar quantities of 2-aminoadenine and thymine and of guanine and cytosine.

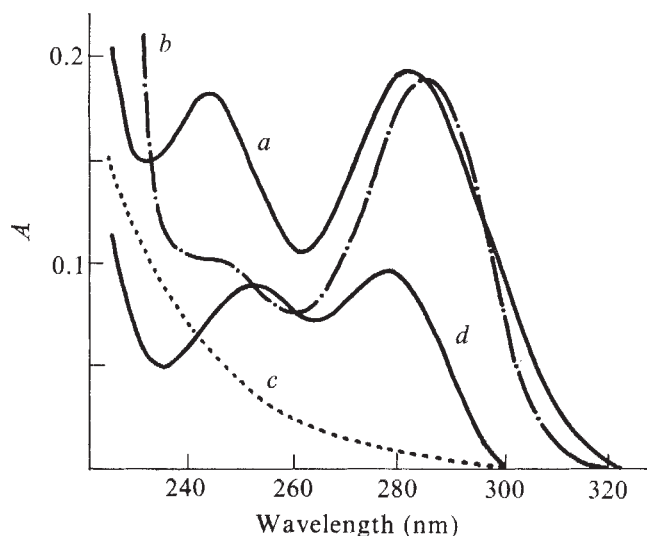
The presence of the additional amino group in 2-aminoadenine as compared to adenine permits formation of the three hydrogen bonds in 2-aminoadenine–thymine base pair¹⁶. This seems to account for the higher T_m (85.6 °C, in 0.1 × SSC) for 2-aminoadenine containing S-2L DNA as compared with usual adenine-containing DNA (T_m = 82.0 °C) which is equivalent to

the former in GC content. According to the T_m value, S-2L DNA should contain 77.4 mol % of GC, since from the ρ value (1.731 g cm⁻³) the GC content should be 72.5 mol %^{17,18}; in fact the DNA contains only 68.7 mol % of GC.

The anomalous CD spectrum of S-2L DNA is similar to the CD spectrum of polyribodiaminopurine–poly U complex¹⁹, the difference being only in the sign of the bands at 290 nm. Dichroism measurements may therefore prove useful for the possible detection of 2-aminoadenine in other viral nucleic acids.

In contrast to other viruses containing novel bases substituting for ordinary pyrimidines^{3–7}, S-2L phage seems to be the first case in which one of the ordinary purine bases in DNA is completely substituted with a modified base. It seems likely that 2-aminoadenine incorporates into S-2L phage DNA as a ready-made nucleotide, rather than forming at the polynucleotide level as a result of amination of adenine residues in DNA. In cells of virus-infected blue-green algae new paths for the synthesis of the adenine nucleotide derivative may therefore exist: a special system for regulation and discrimination between the replication of host and unusual 2-aminoadenine-containing viral DNA would be needed.

Fig. 3 Ultraviolet spectra of the base and its deoxyribonucleoside isolated from S-2L phage DNA. *a*, Base in 0.1 M HCl; *b*, base in 0.1 M KOH; *c*, base after bromination; *d*, deoxyribonucleoside in 0.05 M K-phosphate buffer, pH 6.7.



M. D. KIRNOS
I. Y. KHUDYAKOV
N. I. ALEXANDRUSHKINA
B. F. VANYUSHIN

*A. N. Belozersky Laboratory of Molecular Biology
and Bioorganic Chemistry,
Lomonosov Moscow State University,
Moscow and Biological Institute,
Zhdanov Leningrad State University,
Leningrad, USSR*

Received 23 May; accepted 6 September 1977.

1. Padan, E. & Shilo, M. *Bact. Rev.* **37**, 343–370 (1973).
2. Adolph, K. W. & Haselkorn, R. *Virology* **54**, 230–236 (1973).
3. Wyatt, G. R. & Cohen, S. S. *Biochem. J.* **55**, 774–782 (1953).
4. Kallen, R. G., Simon, M. & Marmur, J. *J. molec. Biol.* **5**, 248–250 (1962).
5. Takahashi, I. & Marmur, J. *Nature* **197**, 794–795 (1963).
6. Kuo, T. T., Huang, T. C. & Teng, M. H. *J. molec. Biol.* **34**, 373–375 (1968).
7. Vanyushin, B. F., Belyaeva, N. N., Kokurina, N. A., Stelmashuk, V. Y. & Tikhonenko, A. S. *Molekularnaya biologiya* **4**, 724–729 (1970).
8. Brunk, C. F. & Leick, V. *Biochim. biophys. Acta* **179**, 136–144 (1969).
9. Fox, J. A., Booth, S. J. & Martin, F. L. *Virology* **73**, 557–560 (1976).
10. Kleinschmidt, A. K. *Meth. Enzym.* **12B**, 361–379 (1968).
11. Wilkins, M. H. F. *Science* **140**, 941–950 (1963).
12. Freifelder, D. *J. molec. Biol.* **54**, 567–577 (1970).
13. Tomasz, J. *Analyt. Biochem.* **68**, 226–229 (1975).
14. Felix, F., Potter, J. L. & Laskowski, H. *J. biol. Chem.* **235**, 1150–1153 (1960).
15. Davol, J. & Lowy, B. A. *J. Am. chem. Soc.* **73**, 1650–1655 (1951).
16. Cerami, A., Reich, E., Ward, D. C. & Goldberg, I. H. *Proc. natn. acad. Sci. U.S.A.* **57**, 1036–1042 (1967).
17. Marmur, J. & Doty, P. *J. molec. Biol.* **5**, 109–118 (1962).
18. Schildkraut, C. L., Marmur, J. & Doty, P. *J. molec. Biol.* **4**, 430–443 (1962).
19. Howard, F. B., Frazier, J. & Miles, H. T. *Biochemistry* **15**, 3783–3795 (1976).

reviews

Tectonic changes in global terms

D. H. Tarling

The Evolving Continents. By Brian F. Windley. Pp. xviii+385. (Wiley: London and New York, 1977.) £14; \$28.

THIS is, without doubt, one of the best of recent books concerned with the Earth's evolution and is the first which really does attempt to see tectonic changes in terms of the total history of the Earth. At last, the Precambrian has been treated as an integral part of the Earth instead of being confined either to specialist books or to forming the odd paragraph added on to an evolutionary story that is otherwise entirely concerned with the Phanerozoic. It would be ingenuous to fault different parts of the book for lack of sufficient data or for inadequate coverage of general ideas, and any such criticism must be considered secondary to the real achievement of combining an optimum of both information and ideas within less than 400 pages.

The level is essentially for third year undergraduates and upwards and should form a standard advanced text, following an Arthur Holmes introduction to the earth sciences, although some amplification of plate tectonic terminology will be needed—lithosphere is, for example, not defined, but it is certainly useful to have the Burke-Dewey "aulacogen" described fully, as this term is still not in most dictionaries. The lack of definition of lithosphere compared with that for aulacogen, in fact, reflects the essentially geological-geochemical approach throughout, with the geophysics largely restricted to palaeomagnetism. Nonetheless, to complain about this lack of geophysical content is, to some extent, unfair, as sufficient books have appeared recently on geophysical plate tectonics in the Phanerozoic to sate most people's appetites.

Precambrian geology and models comprise half of the book with the Phanerozoic and generalities forming the remaining half. Logically, it starts with the Archaean, with one chapter on high grade terranes; one on low grade; and a third is concerned with models of Archaean crustal evolution. The Proterozoic takes a little longer, with four chapters preceding the summary of crustal evolutionary models for this period. The ability to condense the Archaean better than the Proterozoic

possibly reflects Windley's previous background, but it must be stressed that meaningful generalisations about the evolution of Proterozoic regimes have previously been virtually restricted to the publications of Sutton and Watson.

Naturally, such an initial compilation of information and ideas can be criticised in numerous places, although Windley often plays safe by outlining other people's descriptions without significant critical comment. This is understandable in relation to 'factual' observations, but is less satisfactory where ideas are presented and discussed. Albeit rarely, Windley's views sometimes tend to bias the presentation of other schema. Criticism is always possible on the detail of the coverage—half-a-page on the West Australian Kalgoorlie System is barely sufficient to provide references; and this is virtually all it does. But any real extensions would have doubled the size of the book and probably more than doubled its price.

The second half, dealing with the generally accepted Phanerozoic plate tectonics, is less successful, and it is arguable as to the extent to which it could be both reduced and expanded. There are three chapters on orogenic belts—Caledonian, Hercynian and Alpine—which illustrate this point. The Caledonian chapter could well be reduced, as this is described well elsewhere. The Alpine chapter is so abbreviated as to be virtually notes against which a specialist could view his own area or interest, but they are too skeletal for the undergraduate. The chapter on Hercynian is sufficiently short to reflect the surprising lack of interest by most geologists in this period, but does not adequately represent the multiplicity of views which interested earth scientists have expressed. All three, as usual, also suffer from a parochial 'Laurentian' view, with no consideration of the extension of such systems into, for example, Asia.

The second half also suffers from an apparent desire to insert as many aspects as possible. It is a pity, for example, that the elementary outline of palaeomagnetism is not severely reduced to provide space for a discussion of how isotopic studies actually

indicate the source for volcanics, as this knowledge is virtually assumed in the opening chapters. Similarly, the treatment of palaeoclimates is so elementary as to be almost unnecessary. Also crammed within this half are the chapters on seafloor spreading, cordilleran mountain chains, palaeontological implications, island arc systems and an outline of the breakup of Pangaea. Some of these chapters tend to side-track, so that the breakup of Pangaea seems to be mostly related to the evolution of volcanic systems rather than the actual history of break up and separation which is again essentially in note form.

One of the other new features of the book is an attempt to introduce sections of the characteristic mineralisation associated with specific times and localities. This aspect clearly reflects the general, increased interest in the applicability of academic research to real problems. Windley is to be congratulated on his attempt, but this is, unfortunately, probably the least satisfactory part—even though this must be considered a major achievement. The problem is probably one of space, as most of the mineralisation sections are fairly bald statements of what minerals occur where, but with inadequate discussion of why they occur in those particular places at those particular times. To some extent such origins may still be unknown and would therefore be highly speculative, but it is disappointing to find that possible causative factors are not adequately discussed. The same is even more true for the appraisal of the changing tectonic processes. Drastically different tectonic regimes are described for the Archaean, Proterozoic and Phanerozoic, but the only causative feature which is described as being different is that of radiogenic heat production within the Earth. This change in heat production with time is quantified, but Windley does not offer any considerations for the implication of such changes to, for example, lithospheric thickness and the major, almost catastrophic, changes in the physical behaviour of the lithospheric plates which occur as such changes took place. In other words, there is really no discussion of the actual mechanisms—discussions of con-

vection are mostly restricted to minor Archaean problems. On this basis, the book tends to be overly concerned with the end-products of mantle processes and virtually ignores the mechanisms by which they are derived, and therefore why their operation may have changed with time. Some discussion of the vertical variation in the continental crust is also required and the way in which these differences have evolved should have been discussed. Reducing some of the side-lines in the second half and really considering the changing mechanism of tectonic change would then have been more consistent with the title.

The production is excellent and the illustrations are ample; in fact almost too many, as several could have been usefully combined. This tends to apply also to the text, where Windley's desire to give other people's views possibly

would have been better presented as a clear statement with references. These features tend to make some parts seem rather too much like lecture notes, with slides of published work, than is desirable. Such comments are, however, intended sympathetically, as it would be hard to improve the text without grossly expanding the total length. Nonetheless, although improvements are possible, particularly in considering the mechanism of tectonic change, this book can be unreservedly recommended to final-year and postgraduate students both as a textbook and a source of references for further studies. It must be an essential on any final-year earth science reading list. □

D. H. Tarling is Senior Lecturer in the Department of Geophysics and Planetary Physics at the University of Newcastle-upon-Tyne, UK.

Dynamic systems

Synergetics: An Introduction. Non-Equilibrium Phase Transitions and Self-Organisation in Physics, Chemistry and Biology. By H. Haken. Pp. xii+325. (Springer: Berlin and New York, 1977.) DM72; \$31.70.

FOR some time now, it has been the main concern of the author to popularise the subject he has chosen to call "synergetics". To most readers, however, this term is less familiar than the subject itself—the theory of phase transitions in co-operative dynamic systems, applied to examples ranging from physics to sociology. Part of the material presented in this book is already contained in several review articles and conference proceedings, but now H. Haken presents in this volume a detailed and comprehensive account of the basic principles and applications of this "interdisciplinary field of research".

The term "synergetics" was prompted by the variety of dynamic systems stemming from physics, chemistry, biology and sociology which show striking similarities at their points of instability. As the book is meant to be a student textbook, the author provides in the first few chapters the mathematical prerequisites. After some introductory sections on probability and information theory, the reader becomes acquainted with stochastic processes (Brownian motion, random walks and their description by Master equations) and deterministic motion (here, the concepts of stability, critical points, hard and soft mode excitations, limit cycles, and so on, are introduced).

The connection between deterministic and stochastic processes (or chance and necessity) is provided by the Fokker-Planck equation approach, which is discussed extensively and used to introduce the Landau theory of phase transitions. Combining all these concepts, the author then explains in a rather formal and abstract way the generation of order by self-organisation. Then, separately, these formal considerations are applied to various examples. Starting from physics (laser and hydrodynamic instabilities), the reader is led to nonlinear chemical reactions (with their showpieces, Brusselator and Oregonator) and biological systems (models for ecology, evolution and morphogenesis); and, finally, a model for interacting social groups is presented.

This book certainly serves its purpose as an introductory student text, as it requires little previous knowledge. Besides tradition (synergetics is not part of a usual university curriculum), the only obstacle to its wide distribution among students may be its fairly high price.

Experienced readers will certainly profit from the variety of examples presented and the rather detailed reference section, a good guide to the topical literature. It is a pity, however, that some of the important work of several Japanese groups on nonlinear chemical reactions is missing—instead, the author has chosen to cite unpublished work carried out by his own group. But this minor weakness is by far outweighed by many advantages.

R. Schraner

R. Schraner is a Postdoctoral Fellow of the Max-Planck Institute of Biophysical Chemistry, Göttingen, FRG.

Quantum mechanics

The Uncertainty Principle and the Foundations of Quantum Mechanics. Edited by W. C. Price and S. S. Chissick. Pp. 572. (Wiley-Interscience: London and New York, 1977.) £19.50; \$38.

BEING a collection of articles, this book cannot be quite faithful to its title: particularly since it is just one book in a series of books, each designed as a tribute to one of the founders of modern physics. Clearly, such publications are not intended to provide anything like a systematic and critical survey of a subject, especially when the latter is highly structured, as is the case with regard to the foundations of a discipline. But what such books can do is to present significant partial results and to direct the reader's attention toward recent specialised research. This is in fact what is achieved, and in a highly satisfactory way, by this book.

Although I benefited from several of the valuable articles the book contains, let me follow my own taste in selecting just one for review: the contribution by G. Lanz. This paper bears on what is often considered to be the only problem remaining in the foundations of quantum mechanics, namely the problem of measurement. Its description of the difficulties besetting that problem is clear. So is its classification of the two conceptions of quantum mechanics that (barring non-local hidden variables) one may still discuss quantum mechanics as the basic theory of physics or quantum mechanics as a mere theory of microsystems, anchored on a (still to be created) theory of macrosystems. The author considers that the first view should be dismissed because of insuperable difficulties with the problem of measurement. As for the second conception, it obviously requires a theory of macrosystems, and an unambiguous definition of what a macro-system is.

Lanz acknowledges that he has until now no completely definite answers to such questions. He sketches, however (with some profusion of mathematical detail), an approach based on Ludwig's formalism which he believes is promising. It involves a limiting procedure the physical meaning of which is unfortunately not discussed. The general idea is to extract from the N-body theory a part that could be used for a theory of macrosystems; the rest could then be thrown away as physically irrelevant. An interesting but perhaps premature question is whether the future theory will be able to define macrosystems in such a way that at least some finite (and really existing) systems are macroscopic.

Bernard d'Espagnat

Bernard d'Espagnat is Professor of Theoretical Physics at the University of Paris-Orsay.

Outlines of Mercury

The Atlas of Mercury. By Charles A. Cross and Patrick Moore. Pp. 48. (Mitchell Beazley: London; Crown New York, 1977.) £5.95: \$10.

JUDGING by the presence of the word "telescope" in the glossary at the back of this book, Cross and Moore are attempting a relatively popular treatment of their subject. To this end, in addition to the glossary, they provide helpful diagrams to explain anything that might give pause for thought; for example, why Mercury has a day which is two of its years long. This consideration, shown also in the innumerable illustrations and photographs provided, makes the book easy and informative reading for the relatively uninitiated; whereas, on the other hand, its comprehensive nature should mean that it is not without interest for more advanced readers.

The broad scope of the book also provides a neat way of illustrating just how much of our present knowledge of Mercury we owe to that triumph of American technology, Mariner 10, since the whole history of observation (including much duplication) up to the time of Mariner 10, takes up far less of the book than the results gained from that one space-probe. Those who do not have the time to count the relevant number of pages need only compare Chapman's pre-mission map of the planet (p5) with the map drawn by Cross. It ought perhaps to be remarked, however, that Chapman's map is the upper of the two described on p5 and that it is not the latest pre-Mariner map; this distinction belonging to a map drawn by J. B. Murray *et al.* (*Icarus*, 17, 576-584; 1972). Cross's excellent maps, which form the basis of this book, are centred on the useful subsolar point, and the accompanying photographs give an accurate idea of the planet's surface, perhaps the only omission being a high resolution picture of the Hilly and Lineated Terrain. Also, some mention of the processing these pictures have undergone might have been interesting.

The text accompanying the pictures, however, is the weakest part of the book. There was just not enough geological nitty-gritty to maintain my interest in the endless succession of craters; and calling them circular "enclosures", "formations" or "walled plains" did not help. It could be said that this is the criticism of a geologist, and rather unfair, since the book is an elementary one wherein the authors are clearly trying to keep things simple, but the approach of describing things by quadrants and vast numbers of photographs rather than subjects has provided gaps for the astronomy-orient-

ated authors to stray into, and go astray in, the geological field. For example, on p24, it is argued that, because of the lower number of scarps and the "various younger craters and basins", the area covered by the South-West Quadrant is younger on the whole than the South-East Quadrant. In fact, it is the oldest features which indicate the age of the terrain and in terms of age and distribution of craters, the South-West Quadrant could be older, if anything, than the South-East, but whether it is or not, the statement would have been unnecessary in the first place if something more solid had not been cut out by the simplification process.

A key to the age of the surface is the nature of the Intercrater Plains, which on p18 are correctly described as the regions between the main craters (assuming there is nothing else there, such as Smooth Plains). In the next paragraph, however, it is stated that there is an "obvious" resemblance between this intercrater unit and the highlands of the Moon, which for the most part are nothing but craters. The original description (Trask and Guest, *J. Geophys. Res.*, 80, (17), 2461-2477; 1975) is quite clear: "Intercrater Plains (unit) has a closely similar analog on the Moon in the Pre-Imbrian Plains of Wilhelms and McCauley (*Miscellaneous Geological Investigation Map 1-703*, US Geological Survey; 1971)". That is to say, the Intercrater Plains are similar morphologically to terrain situated between the main craters of a relatively sparsely cratered unit within the southern highlands of the Moon. The lunar unit is of much smaller extent and for this reason, among others, is not thought to originate in the same way.

On p30, a prominent ray is said to be "clearly a northward extension of Heemskerck Rupes". If this is so, then it is an exciting discovery linking what has hitherto been considered as a superficial unit, thrown out by the impact of a meteorite, to a feature interpreted as a very large compressional fault scarp in the planet's crust. I am not excited. On p35 Rupes Zeehan is said to be formed from wall remnants of pre-Caloris craters. Certainly, it seems to be influenced by them, but faces the wrong way for the authors' description to be correct.

On p21, it is indicated that the large scarp in the crater Po Ya "seems to be" a lava front rather than a compressional feature. There are good examples of small scarps apparently confined to the Smooth Plains (not present in Po Ya) on some crater floors which some think might be the fronts of highly viscous lava flows (c.f. highly fluid mare basalts). The Po Ya scarp is not one of these. This error is an example of the heavy volcanic bias present in the book. Thus, it is thoroughly misleading of the authors to suggest that there is any controversy over

the cause of the craters on the Moon (and therefore Mercury), even if they do plump for a mixture of both impact and volcanic hypotheses. The vast majority of craters on the Moon, Mars and Mercury were caused by the impact of impinging space debris and because they are morphologically different from the volcanic features also present on the Moon and Mars (but not yet pointed out with any confidence on Mercury), there is no argument as to their origin.

If all this is not entirely compatible with the implication in the proud claim on the cover that this is "the most concise and accurate account of the planet available" (Sir Bernard Lovell), it is also true that, although the mistakes tend to mislead or confuse as to certain pictures or features, they do not seriously misrepresent the fundamental processes which have shaped the planet's surface. In addition, apart from the volcanic bias, the authors do strive throughout to maintain an admirable scientific restraint. For example, the attractive theory that Mercury was once a satellite of Venus is quite properly described as "still highly speculative". On the whole then, this book can be commended to the readers for whom it was written. **W. P. O'Donnell**

W. P. O'Donnell is a research student in planetology at University College, London, UK.

BOOKS

ON PURE AND APPLIED SCIENCE

Books reviewed or mentioned in this journal are available from stock.

Catalogues on application.

Please state interests.

SCIENTIFIC LIBRARY

ANNUAL SUBSCRIPTION from £7.00

Reduced rates for multiple subscriptions

Available in U.K. only

Prospectus free on request

H.K. LEWIS & Co. Ltd.

LONDON: 136 GOWER STREET,
WC1E 6BS

Telephone: 01-387 4282

Satellite technology

The Versatile Satellite. By Richard W. Porter. Pp. +173 (Oxford University: Oxford and London, 1977.) £4.95.

THE author sets himself the task of explaining "what a satellite is, how it is put into orbit (and why it does not promptly fall down again), what special considerations enter into its design, and especially what it is 'good for' and why". Since all of this is attempted in 173 pages, it is not surprising that the reviewer, an environmental scientist, found the treatment of environmental applications was rather superficial. Nevertheless, a great deal of technical information is conveyed in a clear and concise manner.

The book is well structured, beginning with the history of launching the satellites and their evolution from German and Russian rocketry. This is followed by a useful explanation of orbits and energy requirements for different types of satellite configurations. In a discussion of design requirements, emphasis is placed on the special circumstances of lack of atmosphere, absence of gravity, ionosphere, thermal environment, electromagnetic radiation, geomagnetic field, charged particle environment, and micro-meteoroids. The author praises the satellite designer for overcoming this unfriendly satellite environment within limitations of weight and cost. However, the cost-benefit aspects of the satellite programmes are not discussed in depth in this book.

The environmental scientist should have particular interest in those chapters explaining "what it is good for and why". These discuss applications in the fields of communications; weather forecasting and monitoring; navigation; surveying the oceans and land; monitoring flora, fauna and fish; and spaceborne laboratory experiments.

The impressive growth of satellite communications is well documented and the characteristics of Intelsat satellites are described in some detail. The technologist and engineer may revel in the details of stabilisation, antennas and electronic systems, but all readers will flinch in the face of space jargon — for example, SPADE — "a super-acronym for single channel per carrier pulse code modulation multiple access demand assignment equipment". Even Mr Porter is moved to comment that "the jargon gets worse and worse but really there is no way to avoid it". Although agreeing with this, one wishes a glossary of acronyms had been included so that one does not have to hunt through pages to find the explanation of a particular term.

The section of the book dealing with the surveying of the oceans and the

land gives a good technical description of the "Earth's Resources Technology Satellites. Some of the illustrations used to demonstrate different spectral responses are disappointing, especially Fig. 7.6. Also one notes that the concept of the spectral signature, first mentioned on p98, does not receive further explanation until p127. Likewise, problems affecting the spatial extension of supervised and unsupervised classifications of digital data are not discussed. As a result, the casual reader may be left with the false impression that recognition of crop types is a simple matter.

The informed reader will feel that the present capability of LANDSAT in respect of mapping soils and land use is somewhat overstated. One may note that recent pronouncements by NASA officials are more cautious. For example, J. Morrison, (*NASA Earth Resources Survey Program: Problems and Prospects*, European Space Agency Symposium on Remote Sensing of Earth from Space, Strasbourg, 1977) writes "I do not think we will be able to do this (that is, the Worldwide Crop Information System) successfully until about the 1984 period. The reason is that the present LANDSAT satellites simply do not have the resolution necessary for providing worldwide crop data, particularly in small field situations". Likewise, the assertion that plant disease can often be detected in satellite images before it is apparent to the casual observer on the ground was not borne out by the Corn Blight Watch Experiment (R. B. MacDonald *et al.*, *Results of the 1971 Corn Blight Watch Experiment*, Proc. 8th Int. Symp. Remote Sensing of the Environ-

ment, 1, 157, Ann Arbor, Michigan; 1973). It seems likely that such achievements must await the launch of LANDSAT-D and the Thematic Mapper.

Perhaps the most interesting part of this book is the chapter entitled "Laboratory for Science; An Observatory for Astronomy". This presents the case for space experiments and space observations in a succinct and effective manner. Subsequent sections dealing with the manned satellites and the Shuttle programme are also useful. One feels, however, that the role of the European Space Agency in the Spacelab programme could have been given greater prominence. In contrast to the space race between the US and the Soviet Union, the Spacelab programme provides a welcome example of international collaboration. Further discussion of Spacelab would also have revealed some of the initial limitations of the Shuttle in respect of mission duration and maximum inclination of orbit when launched from the Eastern Test Range.

In the last paragraph, Mr Porter notes that the future depends on the use that is made of satellite technology. It is increasingly apparent that there are considerable problems of technology transfer to be overcome before satellite information will be used to its full potential. This book provides an interesting glimpse of the technologist's world and will help the potential user to understand some of the engineering constraints on satellite systems.

L. F. Curtis

L. F. Curtis is Reader in Geography at the University of Bristol, UK.

Parasitic Protozoa

Parasitic Protozoa. Vol. 4: *Babesia, Theileria, Myxosporidia, Microsporidia, Bartonellaceae, Anaplasmatidae, Ehrlichia, and Pneumocystis.* Edited by J. P. Krier. Pp. xv + 386 (Academic: New York and London, 1977.) \$32.50; £23.05.

I HAVE already commented on the fact that the scope of this series encompasses more than the Protozoa (*Nature*, 268, 773; 1977); and favourably, as the treatment of parasitic Protozoa along with other parasitic organisms of similar basic behaviour, multiplying in the vertebrate host, is more likely to be conducive to exchange of fruitful concepts than the traditional association of Protozoa with the helminths, organisms of a basically different mode of life. The present volume is an illustration of this useful format, presenting ten chapters dealing respectively with *Babesia* in domestic animals and in man and other hosts; with *Theileria*, Myxosporidia, Microsporidia,

Bartonella with *Grahamella*, *Anaplasma*, *Aegyptianella*; with *Eperythrozoon* and *Haemobartonella*, Ehrlichiae and *Pneumocystis*. The volume deals with a range of organisms all of which are important in relation to disease in man or his domestic animals, interpreting the latter so as to include his cultivated fish and insects.

The reviews are excellent up-to-date summaries of their subjects, many of which are diffusely scattered in the literature. They will be popular as systematic first readings or as reference guides to the 'state of play' of their respective subjects. The book is well presented and indexed, with a few line diagrams and half-tones; the last do not receive special treatment as regards the paper on which they are printed and so are adequate rather than excellent.

W. H. R. Lumsden

W. H. R. Lumsden is Professor of Medical Protozoology at the London School of Hygiene and Tropical Medicine, London, UK.